Improving estimates of children living with HIV from the Spectrum AIDS Impact Model

Mary Mahy\textsuperscript{a}, Martina Penazzato\textsuperscript{b}, Andrea Ciaranello\textsuperscript{c}, Lynne Mofenson\textsuperscript{d}, Constantin T. Yianoutsos\textsuperscript{e}, Mary-Ann Davies\textsuperscript{f} and John Stover\textsuperscript{g}

Objective: Estimated numbers of children living with HIV determine programmatic and treatment needs. We explain the changes made to the UNAIDS estimates between 2015 and 2016, and describe the challenges around these estimates.

Methods: Estimates of children newly infected, living with HIV, and dying of AIDS are developed by country teams using Spectrum software. Spectrum files are available for 160 countries, which represent 98% of the global population. In 2016, the methods were updated to reflect the latest evidence on mother-to-child HIV transmission and improved assumptions on the age children initiate antiretroviral therapy. We report updated results using the 2016 model and validate these estimates against mother-to-child transmission rates and HIV prevalence from population-based surveys for the survey year.

Results: The revised 2016 model estimates 27% fewer children living with HIV in 2014 than the 2015 model, primarily due to changes in the probability of mother-to-child transmission among women with incident HIV during pregnancy. The revised estimates were consistent with population-based surveys of HIV transmission and HIV prevalence among children aged 5–9 years, but were lower than surveys among children aged 10–14 years.

Conclusions: The revised 2016 model is an improvement on previous models. Paediatric HIV models will continue to evolve as further improvements are made to the assumptions. Commodities forecasting and programme planning rely on these estimates, and increasing accuracy will be critical to enable effective scale-up and optimal use of resources. Efforts are needed to improve empirical measures of HIV prevalence, incidence, and mortality among children.

Introduction

Globally, there has been widespread scale-up of services for pregnant women living with HIV to improve their own health and prevent HIV transmission to their children. As a result of this scale-up, it is estimated that globally, the annual number of new child HIV infections (among children <15 years) dropped by 50% between 2010 and 2015 \cite{1}. In 2015, an estimated 77% (69–86%) of pregnant women living with HIV received antiretroviral prophylaxis to prevent mother-to-child transmission (MTCT) or lifelong antiretroviral therapy (ART) \cite{1}.
The 2015 WHO guidelines recommend starting all individuals living with HIV on ART [2], but only an estimated 49% (42–55%) of children living with HIV (CLHIV) received ART in 2015 [1]. Efforts are needed to expand access to paediatric treatment and to accurately estimate the number of CLHIV at national and global levels to support paediatric ART programme planning [3]. Accurate estimates of the extent of the paediatric HIV epidemic are also necessary to forecast demand for HIV commodities and age-appropriate antiretroviral formulations; these estimates in turn inform the quantity of paediatric antiretroviral drugs that need to be manufactured and procured to ensure a reliable supply of paediatric antiretrovirals, and also to better target and prioritize the development of suitable paediatric antiretroviral drug formulations [3]. The accuracy of estimates is also becoming more important as the numbers of vertically acquired infections decrease.

In 2015, an estimated 1.8 million (1.5–2.0 million) children aged 0–14 years were living with HIV, an estimated 150,000 (110,000–190,000) children were newly infected, and 110,000 (84,000–130,000) children died of AIDS-related causes [1]. These global estimates are based on mathematical models, because robust empirical data on numbers of CLHIV are limited, and can never be completely comprehensive. Data are particularly limited about the total number of infants newly acquiring HIV infection, and about rates of survival among CLHIV, particularly for children infected postnatally (during breastfeeding).

Spectrum is a well established model that has generated robust global AIDS estimates since 1999 [4]. The model is primarily driven by adult HIV surveillance data, whereas demographic and programme data are used to further calculate the HIV epidemic among children. Since 2008, the UNAIDS (Joint United Nations Programme on HIV/AIDS) Reference Group on Estimates, Modelling and Projections has held five consultations among expert groups to address different challenges related to child estimates (including reviewing the structure of the child model, MTCT probabilities and survival on ART [5]; estimating ART needs among children [6]; re-evaluating the estimated probabilities of MTCT [7] and, convened with WHO, reviewing the model structure and assumptions) [8]. During each of these consultations, important model modifications changed the estimated numbers of new child HIV infections, CLHIV, and AIDS-related deaths among children (see Annex 1, http://links.lww.com/QAD/A999 for a description of previous adjustments to the UNAIDS child HIV estimates).

In this study, we summarize key changes implemented in the Spectrum child model between 2015 and 2016, the research that informed these modifications, and compare the modelled estimates to the best available empirical data. We highlight limitations and persisting evidence gaps to be addressed in future model revisions as part of the ongoing effort to obtain more accurate global and national paediatric HIV estimates.

Methods

Structure and parameters

The Spectrum AIDS impact module (AIM) estimates children newly infected with HIV each year, progresses them through a series of health states over time and transitions surviving children to the adult model at 15 years (see Fig. 1).

The AIM module calculates HIV prevalence among adults aged 15–49 years from prevalence trends, and is described elsewhere [9]. HIV prevalence among pregnant women is derived from age-specific HIV prevalence and age-specific fertility rates among all women, with adjustments for the effects of HIV infection on fertility [10]. The child model only calculates new HIV infections due to MTCT and assumes no new infections from other sources other than MTCT among children aged less than 15 years. Rates of transmission depend on maternal CD4+ cell count for women who did not receive any antiretroviral-based prophylaxis or treatment, and by type of regimen for those who did receive antiretrovirals. For women receiving lifelong ART (excluding women receiving a temporary antiretroviral-based prophylaxis), the duration of treatment during peripartum is categorized as started before conception, more than 4 weeks before delivery or less than 4 weeks before delivery. Women on lifelong ART are assumed to remain on therapy throughout breastfeeding, with a default dropout rate of approximately 2% per month. If the country team has data on dropout among breastfeeding women, they can enter national dropout values.

Transmission probabilities

In 2012, a review of existing literature identified 32 studies to determine the probability of transmission for women receiving no antiretroviral and those receiving different antiretroviral regimens [11]. In 2015, an additional 48 studies were added to the 2012 analysis (see Annex 2, http://links.lww.com/QAD/A999 for a full list of studies used for this analysis and a further explanation of the regimen categories). The updated transmission probabilities were included in the 2016 Spectrum model.

In the Spectrum model, transmission may occur in utero or intrapartum (peripartum), or through breastfeeding (postnatal). The model calculates the numbers of new infections separately for peripartum HIV transmission and postnatal transmission.

Figure 2 presents the changes to the probability of MTCT between the 2015 and the 2016 round of HIV estimates.
The peripartum values describe the probability of transmission during pregnancy and delivery, and reflect any lack of adherence that existed in the study settings (for some cases, the adherence from the studies will be better than adherence in the general population). The postpartum values are monthly probabilities and are multiplied by the approximate duration of breastfeeding in months. Transmission due to an incident infection during breastfeeding is modelled as a one-time risk and is thus not included in Fig. 2. The transmission probability for an incident infection during breastfeeding is 27% (compared to 28% in the 2015 model).

The largest adjustment occurred in the assumption about transmission among women with an incident infection during pregnancy (previously 30%, revised to 18%). The probability of 30% came primarily from one study in South Africa [12]. The more recent review of available data revised the probability after including data from additional studies including two that found transmission risk less than 15% among maternal incident infections during peripartum.

Disease progression among children not receiving antiretroviral therapy
After infection, the model progresses children separately depending on the age at which infection occurs (at 0–6 months, 7–12 months or greater than 12 months after birth). The median survival without treatment is less than 2 years for children infected perinatally, but increases to approximately 6 years for those infected through breastfeeding at 0–6 months, 11 years for those infected at 7–12 months and 14 years for those infected after 12 months [13,14] (see Annex 3, http://links.lww.com/QAD/A999 for further description).

The model transitions CLHIV through a series of CD4+ categories in the absence of ART. Children under 5 years are followed through seven CD4+ cell percentage categories. Newly infected children are allocated to CD4+ cell percentage categories based on data from the HIV Pediatric Prognostic Markers Collaborative Study (HPPMCS) and Cross Continents Collaboration for Kids [6,15] (see Annex 3, Table 3.1, http://links.lww.com/QAD/A999). The distribution of new infections by CD4+ cell percentage category and the progression rates are assumed to be independent of timing of transmission (peripartum versus postnatal), but are dependent on age.

Each year, the HIV-infected paediatric population ages by 1 year and is subject to non-AIDS mortality, HIV-related mortality and progression to a lower CD4+ category. At age 5 years, children transition from CD4+ cell percentage categories to CD4+ cell absolute count categories. The transition is based on the CD4+ cell distributions reported by HPPMCS for children at age 5 years (see Annex 3, Table 3.2, http://links.lww.com/QAD/A999). The HPPMCS data are available for children aged 0 through 10 years. The data for children aged 10 were assumed to apply to children...
Mortality rates and progression rates by CD4\(^+\) cell category are specified by age group and age at infection. These rates have been derived by fitting a cohort model to paediatric survival patterns [16]. Mortality rates were constrained to increase exponentially with lower CD4\(^+\) cell category (see Annex 3, Table 3.3, http://links.lww.com/QAD/A999 for the resulting mortality rates).

Children initiating antiretroviral therapy

The number of children (0–14 years) receiving ART in each calendar year is based on programme data entered in the model by country teams. The Spectrum model calculates the number of children who must start ART each year to match programme data, while taking into account mortality of those on treatment. In the 2015 model, the age distribution of children starting on ART was assumed to imitate the age distribution of when children became eligible for ART under national ART eligibility criteria. In the 2016 model, the age distribution of initiating ART was adjusted to reflect the age distributions from patient data collected through the International Epidemiological Databases to Evaluate AIDS (IeDEA) Consortium (www.iedea.org).

Data on children starting ART up to 15 years were used from six IeDEA regions: Asia-Pacific; Caribbean, Central and Latin America; and West, South, Central and East Africa regions. In most cases, data were provided for the year 2001 onward. All children who started ART for the first time at the sites participating in the IeDEA network were included [17–19]. Children who were transferred aged 11–14 years due to a lack of alternative information on CD4\(^+\) progression without ART.

Mortality rates and progression rates by CD4\(^+\) cell category are specified by age group and age at infection. These rates have been derived by fitting a cohort model to paediatric survival patterns [16]. Mortality rates were constrained to increase exponentially with lower CD4\(^+\) cell category (see Annex 3, Table 3.3, http://links.lww.com/QAD/A999 for the resulting mortality rates).

Children initiating antiretroviral therapy

The number of children (0–14 years) receiving ART in each calendar year is based on programme data entered in the model by country teams. The Spectrum model calculates the number of children who must start ART each year to match programme data, while taking into account mortality of those on treatment. In the 2015 model, the age distribution of children starting on ART was assumed to imitate the age distribution of when children became eligible for ART under national ART eligibility criteria. In the 2016 model, the age distribution of initiating ART was adjusted to reflect the age distributions from patient data collected through the International Epidemiological Databases to Evaluate AIDS (IeDEA) Consortium (www.iedea.org).

Data on children starting ART up to 15 years were used from six IeDEA regions: Asia-Pacific; Caribbean, Central and Latin America; and West, South, Central and East Africa regions. In most cases, data were provided for the year 2001 onward. All children who started ART for the first time at the sites participating in the IeDEA network were included [17–19]. Children who were transferred aged 11–14 years due to a lack of alternative information on CD4\(^+\) progression without ART.

Mortality rates and progression rates by CD4\(^+\) cell category are specified by age group and age at infection. These rates have been derived by fitting a cohort model to paediatric survival patterns [16]. Mortality rates were constrained to increase exponentially with lower CD4\(^+\) cell category (see Annex 3, Table 3.3, http://links.lww.com/QAD/A999 for the resulting mortality rates).

Children initiating antiretroviral therapy

The number of children (0–14 years) receiving ART in each calendar year is based on programme data entered in the model by country teams. The Spectrum model calculates the number of children who must start ART each year to match programme data, while taking into account mortality of those on treatment. In the 2015 model, the age distribution of children starting on ART was assumed to imitate the age distribution of when children became eligible for ART under national ART eligibility criteria. In the 2016 model, the age distribution of initiating ART was adjusted to reflect the age distributions from patient data collected through the International Epidemiological Databases to Evaluate AIDS (IeDEA) Consortium (www.iedea.org).

Data on children starting ART up to 15 years were used from six IeDEA regions: Asia-Pacific; Caribbean, Central and Latin America; and West, South, Central and East Africa regions. In most cases, data were provided for the year 2001 onward. All children who started ART for the first time at the sites participating in the IeDEA network were included [17–19]. Children who were transferred aged 11–14 years due to a lack of alternative information on CD4\(^+\) progression without ART.
from other clinics were excluded because the dates of earliest ART initiation were not reliable in most cases. The IeDEA data were transformed into the age-specific probability that a CLHIV who is not on ART will start ART each year.

The number of CLHIV starting ART in each calendar year and age group is calculated by subtracting the programme reported number of children on ART from the estimated number of CLHIV to estimate the number of CLHIV not receiving ART, multiplying the estimated number of CLHIV not on ART in each age group by the empirical probability of starting treatment derived from the IeDEA data, and adjusting the age distribution to the correct number of children starting ART in that year. This approach preserves the age pattern of starting ART while adjusting the total number starting ART to match programme records (see Annex 4, http://links.lww.com/QAD/A999 for information on the validation studies).

Survival of children receiving antiretroviral therapy
The survival of children on ART is also based on data available from IeDEA. Data were obtained from over 48 000 children aged 0–10 years at the start of ART receiving care in facilities participating in the IeDEA collaboration. Programmatic data spanned up to 15 years and were collected from the four regions in Africa, and the Asia-Pacific and the Caribbean, Central and South America regions. Mortality rates at 6, 12, 24 and more than 24 months after ART initiation were produced and trends in mortality over time were assessed. Mortality rates were adjusted for potential biases due to under-reporting of mortality among children disengaged from care using data obtained from patient-tracking studies or linkage with vital registries when available [20]. Annual mortality rates were calculated by sex, age, CD4⁺ cell percentage (for children aged 0–5 years at the start of ART) or absolute count (for children aged 6–10 years at the start of ART), and region. Data were not disaggregated by sex, as sex was not found to be a significant risk factor for mortality.

Model output (estimates)
Spectrum generates a number of results, including numbers of new HIV infections, people living with HIV, and AIDS-related deaths and their respective rates by age and sex. With every version of the model, revised estimates are generated for every calendar year, from 1970 through the last year of available input data. Surveillance and programme data for 2015 were not available to include in the models developed by countries in 2015 (the 2015 model), and thus the final calendar year with estimates from the 2015 model is 2014. For the 2016 model, the last year with available estimates is 2015. Although it is possible to project the HIV epidemic with Spectrum, however, that discussion is not included in this study.

Model validation
There are little empirical data to describe HIV among children. We validated the 2016 Spectrum model revisions by comparing MTCT risk at 4–6 weeks and at the conclusion of breastfeeding using population-level impact studies; and HIV prevalence among children aged 5–9 years and 10–14 years using prevalence measured in population-based surveys. The studies were limited to nationally representative, population-level, empirical studies that reflect the impact of HIV on children. A PMTCT impact study from Malawi was excluded as it only measured outcomes from immunization clinics in four districts [21] (see Annex 5, http://links.lww.com/QAD/A999 for information on the validation studies).

Results
Comparison of Spectrum estimates for calendar year 2014
In 2016, Spectrum files were available for 160 countries representing 98% of the global population. The comparison of the estimates from the 2015 and 2016 models are shown in Fig. 3. For this comparison, we describe the results for the calendar year 2014 because 2014 is the last year for which the 2015 model produces estimates. The number of new infections in 2014 dropped from 220 000 (190 000–260 000) in the 2015 Spectrum model to 160 000 (130 000–220 000) in the 2016 model – a 27% reduction. The most important reductions in new child infections were during years in which HIV incidence among women was estimated to be highest – in the late 1990s and early 2000.

The revised assumptions about ART initiation in the 2016 model increased the estimated median age of starting ART (see Annex 6, http://links.lww.com/QAD/A999), and, in turn, a greater proportion of CLHIV died before starting on ART. However, the lower overall numbers of CLHIV resulting from lower MTCT probabilities meant that the absolute number of AIDS-related deaths among children was lower than in the 2015 model. The estimated number of children dying of AIDS-related causes in calendar year 2014 was 150 000 (140 000–170 000) in the 2015 model and 120 000 (95 000–150 000) in the 2016 model.

Validation of Spectrum projections
Mother-to-child-transmission risk
Compared to two studies of population-level MTCT rates, the 2016 modelled estimates are within the confidence intervals (CIs), or comparable, to the study results [22,23]. The modelled transmission rates for South Africa at age 6 weeks in the calendar year 2011 was 4.0%, whereas the transmission rate from the South African Impact Assessment was 3.5% (2.9–4.1%) [22] (see Fig. 4).
In Zimbabwe, the study results are for children aged 9–18 months. Preliminary results from the most recent Zimbabwe Demographic and Health Survey suggest that in 2015, 23% of children (regardless of maternal HIV status) were still breastfeeding at age 18 months [24], suggesting that the study results could potentially underestimate the final MTCT rate. In calendar year 2012, modelled estimates of MTCT risk at the end of breastfeeding were the same as the transmission rate from the Zimbabwe PMTCT study: 10.0% (95% CI 7.5–12.6%). In calendar year 2014, the model estimated final transmission rate was 6.8%, whereas the study...
transmission rate at 9–18 months was 4.8% (95% CI 3.4–6.3%).

**HIV prevalence**

Figure 5 compares paediatric HIV prevalence from national surveys to modelled estimates, using the survey year as the year of comparison. Among the 5–9-year age group, the HIV prevalence rate from the 2016 model falls within the CIs of the survey data for three countries with available CIs. The results fall outside of the CIs for Swaziland. For the age group 10–14 years, the prevalence rates from the 2016 model are within the CIs of the survey data for Kenya and marginally outside of the intervals for South Africa. However, the modelled results are lower than the survey data, and outside of the survey CI in Mozambique and Swaziland. For Botswana, for which no CIs were available, the 2016 estimates were considerably lower in both age groups than the survey HIV prevalence.

**Discussion**

In 2016, UNAIDS introduced significant changes to the models that estimate HIV among children aged 0–14 years. Our review has three main findings. First, the 2016 models estimate fewer new child HIV infections since the start of the HIV epidemic. As a result, the numbers of CLHIV and AIDS-related deaths were also considerably lower. Comparing results for calendar year 2014, the estimated CLHIV is 27% lower in the 2016 model due to model refinements. This adjustment is important for our global understanding of paediatric HIV as it implies that fewer children were in need of ART, and

---

**Fig. 4.** Mother-to-child transmission rates based on studies from South Africa [15] and Zimbabwe [16] compared to the 2016 model results. In the South African study, the MTCT rate was measured at 8 weeks, whereas the modelled estimate is for approximately 6 weeks. In the Zimbabwe study, the MTCT rate was measured at 9–18 months, whereas the final modelled transmission rate is estimated at 36 months. MTCT, mother-to-child transmission.
thus ART coverage among children is similar to ART coverage among adults [49% (42–55%) among children versus 46% (43–50%) among adults, globally] [1].

Second, the results from the 2016 models are partially consistent with empirical data from nationally representative population-level studies of MTCT risk in two countries. A further study of MTCT by Silohol et al. [25] compared a sub-national estimate from Spectrum to empirical data from a demographic and surveillance site and found consistency with the empirical data. However, there are mixed results when comparing the estimates with nationally representative HIV prevalence for the age group 5–9 years and 10–14 years. The estimates appear to underestimate prevalence among children aged 10–14 years. The number of CLHIV at older ages might be underestimated, and more children might need ART than the number currently projected.

Third, whereas the reduction in new infections was an important driver in the adjusted number of CLHIV, the age at ART initiation also had an important impact on estimating the survival of CLHIV due to the high mortality in the first 2 years of life. This emphasizes the importance of capturing the timing of diagnosis and ART initiation by age group in the models to improve the accuracy of the national and global estimates of CLHIV.

Limitations to the 2016 estimates remain. Data from the IeDEA consortium on initiation of ART or survival on ART might be biased due to the facilities included in the network, where earlier identification and treatment, combined with better retention, may lead to improved clinical outcomes. The default distribution of when children initiate ART, from the IeDEA consortium, was found to be similar to programme data in two countries with good paediatric ART scale-up (South Africa and Tanzania, see Annex 7, http://links.lww.com/QAD/A999 for a comparison of the IeDEA data to national programme data); however, additional comparisons of the age distribution of ART initiation are needed for other countries and other regions. If the IeDEA data include sites with higher-quality services, the current estimates might overestimate survival among CLHIV.
Lack of empirical, representative data on outcomes of children (such as loss-to-follow-up or disengagement from care) is especially challenging. Estimates of new child infections will be erroneous without adherence and viral suppression data among pregnant or breastfeeding women living with HIV. Few programs have conducted an evaluation of MTCT rates after breastfeeding cessation; if treatment adherence wanes over time outside of the clinical trial setting, the long-term risk of HIV transmission through breastfeeding may be underestimated from the clinical trial data.

The programme data entered into the Spectrum files to produce estimates includes the number of women who received different antiretroviral regimens. Over time, as the antiretroviral regimens for preventing MTCT have changed and the availability of treatment has increased, and countries have struggled to update their registries to capture the required data. The uncertainty bounds in the current models do not reflect the uncertainty of the programme data that are entered into the models. Programme data directly impact the estimated new child infections, CLHIV, ART coverage and AIDS-related deaths.

There is insufficient information of how fertility among women living with HIV changes over time as fertility norms in society change, ART coverage increases and women’s perception of their survival changes. When modelling HIV among children, the number of pregnant women living with HIV is the first calculation. The precision of these numbers directly impacts the accuracy of all subsequent estimates related to CLHIV. Research is ongoing to understand the dynamics of changing fertility for women living with HIV when they are on or off ART [26]; those studies will improve the models in the future.

Additional information to improve the models – including improved PMTCT programme data, data to improve the parameter assumptions or data against which to validate the results – is needed. Even in countries with large surveys, the CIs around paediatric HIV prevalence rates are large, making it difficult to validate the estimates. Collecting HIV prevalence among children in household surveys is only appropriate in countries where adult (15–49 years) female prevalence is 5% or more [27]. For countries with prevalence below 5%, the sample size required to accurately measure prevalence, including determining trends over time, is prohibitively large. In addition, surveys often do not capture the status of children who might have died between birth and data collection.

As global and national estimates are improved, more attention is required to optimize the use of estimates to inform programme scale-up. Current global commodities forecasting and national quantification exercises rely on Spectrum estimates. As ART programmes evolve with innovative testing strategies (multiple testing at different time points for infant diagnosis), broader use of antiretrovirals (such as multidrug postnatal prophylaxis) and increasing use of second-line drugs, new modelling approaches for forecasting will need to be explored to ensure that age-appropriate formulations, and test-kits are delivered when needed in a timely manner.

Similarly, broader programme planning and provision of services is largely based on the estimated number of infants, children and adolescents in need of the different services. As programmes shift toward differentiated care and service delivery tailored to specific ages, accurate estimates will be increasingly important. The potential impact of inaccurate estimates is substantial.

The current global estimates of HIV among children are based on the best data available. They are calculated with more precision than previous models. The improvements in our understanding of when children acquire HIV infection, when they initiate ART and survival of infected children will continue to evolve in the future. As additional empirical data are generated from more robust and routine sources, our ability to validate the modelled estimates on HIV among children will improve. Thus, while the 2016 models produced our best estimates to date, they include substantial uncertainty and are likely to continue to evolve in the future. Efforts to improve these models will be essential to support the global response and reach the goal of an AIDS-free generation.

Acknowledgements

The authors wish to acknowledge the work of the country teams which produce the estimates. M.M., M.P. and A.C. developed the framework of the paper; L.M. conducted the literature review for the MTCT probabilities; C.Y. and J.S. calculated the probabilities of starting on ART; C.Y. and M.D. produced the estimates of survival on ART; J.S. designed and implemented the changes in the software; and M.M. compiled and validated the estimates and the empirical results.

Conflicts of interest

There are no conflicts of interest.

References
