Improving pediatric HIV estimates in Brazil with empirical data

Magdalene K Walters

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Public Health

University of Washington 2022

Committee:

Hmwe Kyu

Alison L Drake

Jeffrey W Eaton

Program Authorized to Offer Degree:

Department of Global Health

University of Washington

Abstract

Improving pediatric HIV estimates in Brazil with empirical data

Magdalene Walters

Chair of the Supervisory Committee:

Hmwe Kyu

Department of Global Health

Despite the availability of methods to prevent vertical transmission of HIV to children, each year there are many pediatric cases of HIV globally. This represents a new generation of people living with HIV, whose odds of surviving until adulthood improve in tandem with treatments for HIV. Estimating burden amongst this age group is not necessarily straightforward; while some countries have Early Infant Diagnosis data available, it is often plagued by underreporting. To combat this, we used Brazil's thorough disease surveillance systems and high quality vital registration to refine maternal prevalence estimates and from this, estimates of mortality and morbidity in children under five. Our analysis found that this approach can be integrated within the Spectrum modeling framework, and that doing so reduces the difference between modeled estimates of disease burden and case report data for children under five. This work provides a platform to continue refining estimates in pediatric HIV burden.

Introduction

In 2020, 187 children under five were diagnosed with HIV in Brazil.[1] As nearly all HIV transmission in the under five age group is vertical, pediatric HIV incidence is driven by the prevalence of HIV among pregnant women and the rate of mother-to-child HIV transmission (MTCT). Burden of HIV in this age group is dependent on incidence and pediatric survival of children with HIV. Estimation of pediatric HIV in Brazil is limited by the lack of Early Infant Diagnosis (EID), insufficient information informing estimates of survival rates for children infected during breastfeeding, and limited information about changes in fertility over time among women living with HIV as attitudes and realities of living with HIV change.[2]

Pediatric HIV estimates are produced by a number of groups. The Joint United Nations Programme on HIV/AIDS (UNAIDS) and the Global Burden of Disease (GBD) study utilize the Spectrum pediatric model. The pediatric model estimates burden due to MTCT from prevalence among pregnant women and rates of transmission calculated in the Spectrum adult model. The broader Spectrum adult model is a disease progression and demographic projection model developed the UNAIDS reference group and implemented by Avenir Health.[3]

Spectrum does not calibrate pediatric incidence to data on HIV infections or deaths from AIDS, and instead uses model assumptions regarding transmission and treatment coverage to simulate disease burden. To utilize observational data of disease burden within the Spectrum framework, UNAIDS and GBD have different ways to combat this limitation. The former requires assumptions surrounding time until case detection and case detection rate provided by countries[4] while the latter utilizes a five year time lag.[5]

Stratified CD4 counts at diagnosis from case reports provide proxy estimates of the lag between HIV infection and diagnosis.[6] This data provides the opportunity to refine input incidence hazard into Spectrum, in turn improving estimates of maternal prevalence, and subsequently improving on HIV burden in children under five. This analysis utilized case reports and deaths reported in vital registration (VR) to triangulate maternal prevalence and improve estimates of pediatric incidence measured against case reports.

Methods

Data sources

The accuracy of Spectrum results can be assessed against VR data about the number of AIDS deaths, for women of childbearing age and children under the age of five. In Brazil, VR data is available for all years between 1990 and 2019.[7] This analysis utilized age, sex stratified case report data; age, sex, and CD4 stratified HIV case report data; VR data; and antiretroviral treatment (ART) data provided by the Brazilian Department of Diseases of Chronic Conditions and Sexually Transmitted Infections.

Data type	Description	Use	Source
Case reports	HIV diagnosis	Informs input incidence hazard	Notifiable Diseases Information System (SINAN)[5]

Case reports, stratified by CD4 count	HIV diagnosis with CD4 count	Refine incidence hazard from case reports by improving assessment of lag between infection and diagnosis	Brazilian government AIDS database
Vital Registration data	Deaths due to HIV stratified by year, age, and sex	Model calibration	Brazil Mortality Information System 1980 – 2019[8]
ART coverage	Percent of people living with HIV on ART, by sex	Used to update ART coverage in efforts to align estimates of HIV mortality with VR	Department of Diseases of Chronic Conditions and Sexually Transmitted Infections, Brazil

Table 1. Description of data sources

Back calculation of incidence

This analysis used the Spectrum pediatric modeling framework. Spectrum relies on an input incidence hazard for the general population aged 15-49 to simulate a disease progression model and create full time series estimates of incidence, prevalence, and HIV-mortality. Pediatric incidence is derived from maternal HIV prevalence and MTCT rates with and without the use of PMTCT treatments.

Case reports were utilized to inform the incidence hazard for 15-49 year olds used by Spectrum to produce pediatric estimates. We utilized CD4 stratified case reports to calculate the distribution of lags between infection and diagnosis. The distribution of cases by CD4 count, probability of progressing between CD4 categories, and probability of surviving one year for each CD4 category (>500, 350-500, 250-249, 100-199, 50-99, <50) were derived from the rates presented in Glaubius et. al.[9] The number of years between infection and diagnosis was capped at ten years, as previous studies have estimated that approximately 80% of all cases are diagnosed within ten years of infection.[10] We assumed that the lag between infection and diagnosis was at least one year, and that in a year an individual could either remain in the current CD4 category, progress to the subsequent category, or die. The probability *p* of being diagnosed *y* years after infection for each CD4 category *c* and age *a* was calculated as:

$$\sum_{y=1}^{10} p_{c,a,y} = \left(\sum_{c=1}^{c-1} o_{c-1} * \left(1 - r_{a,c}^{y}\right) * \left(1 - d_{a,c}\right)^{y}\right) + o_{c} * r_{a,c}^{y} * \left(1 - d_{a,c}\right)^{y}$$

Formula 1

In formula 1, we represent the probability of being in a CD4 category at infection as *o*, the probability of remaining in a CD4 category (*c*) for a number of years *y* stratified by age *a* as *r*, and the probability of survival until age and CD4 before diagnosis as *d*. Additionally, we included the probability of death before HIV diagnosis but following HIV acquisition as shown in formula 2.

$$\sum_{y=1}^{10} p_{c,a,y} = \left(\sum_{c=1}^{c-1} o_{c-1} * \left(1 - r_{a,c}^{y}\right) * d_{a,c}^{\ y}\right) + o_{c} * r_{a,c}^{y} * d_{a,c}^{\ y}$$

Formula 2

These probabilities were summed to determine the stratified distribution of years between infection and diagnosis for a given year and age. The relative distribution of probabilities was applied to the time series of case reports as an estimate of incident cases. The number of incident cases was divided by the susceptible population to calculate the incidence hazard, which was input into Spectrum to create a new time series of age- and sex- specific morbidity and mortality estimates.

Treatment alignment

In Brazil, PMTCT treatment consisted of ART initiation before pregnancy, ART initiation during pregnancy, maternal triple antiretroviral (ARV) prophylaxis (option A), and maternal azidothymidine (AZT) and infant ARV prophylaxis (option B). Using updated results from the case-report informed Spectrum run, option A and option B were scaled to align with the new estimates of births from women with HIV.

Using new general population ART coverage from Brazil's Department of Diseases of Chronic Conditions and Sexually Transmitted Infections (Table 1), we used a linear regression model for all locations in the Latin America and Caribbean region to estimate ART coverage amongst pregnant people based on estimates from the general population.

$$R_1 = \alpha + \beta * year + \varepsilon$$

Formula 3

$$R_2 = \alpha + \beta * year + \varepsilon$$

Formula 4

In the formulas above, R_1 represents the proportion of people in the general population on ART (ART coverage) divided by ART coverage among pregnant people who initiated ART before pregnancy, and β represents the annual change in this proportion. The proportion of ART initiation before pregnancy and after pregnancy is represented by R_2 with β representing the change in this ratio by year. R1 and R2 then used to generate ART coverage in pregnant people before and during birth as input into Spectrum to create final adjusted results.

Results

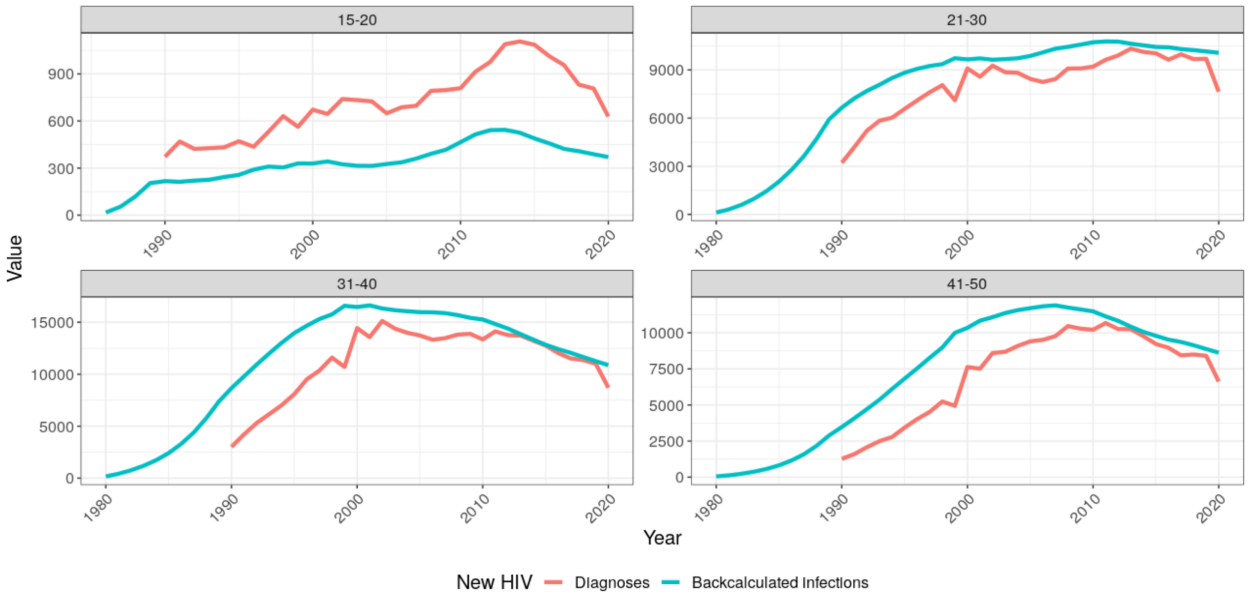

Figure 1. Incidence inference from case reports

Using CD4 stratified case reports, we created a full time series estimate of HIV infections. In the adjusted case reports, the peak number of HIV diagnoses occurred in 2018. The 31 to 40 age group had the highest proportion of cases among all ages with 37.4% of all cases (11,357 diagnosed HIV cases). After back calculating incidence, the highest number of infections occurred in 2012, with the majority of cases occurring in the 31 to 40 year age group (Figure 1). This age group had 14,386 estimated cases in 2012. Over the time series, there were an estimated 29,887 cases of HIV that died before diagnosis, yielding a 93% case detection rate.

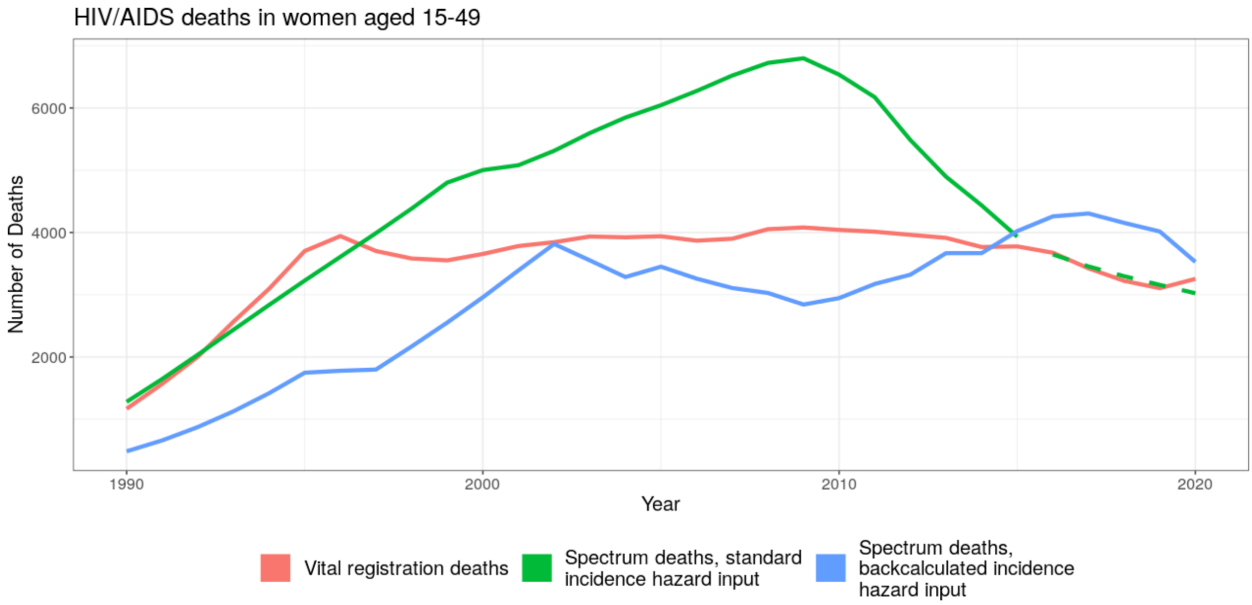

Figure 2. Deaths due to HIV in women aged 15-49 by input. Note: standard incidence hazard post-2015 is extrapolated.

The new incidence hazard input informed by case reports produced estimated deaths that were lower than deaths estimated by Spectrum and deaths present in VR data up until 2015 (Figure 2). The difference between the new modeled deaths and deaths produced with Brazil's Spectrum file reached a peak in 2009, when new modeled deaths were 58.18% lower than deaths from Brazil's Spectrum file. In the same year, the new modeled deaths were 30.33% lower than those present in VR data. The smallest difference occurred in 2015, when all three sources of death estimates (or observation in regards to those from VR data) ranged from 3776.54 to 4023.02.

The updated incidence hazard resulted in lower numbers of births to people living with HIV (henceforth, HIV births) before 1995, but higher between 1995 and 2018 compared to Spectrum results from Brazil's publicly available file. After applying country-provided PMTCT coverage to the new number of HIV births, the number of people on PMTCT options A and B decreased historically. Our estimates of pregnant people on ART before and during pregnancy predicted using formulas 3 and 4 and general population coverage (table 1) resulted in higher ART coverage throughout the time series. There was no difference over time in the ratio of people on ART before pregnancy to the total population, while the ratio of people on ART before pregnancy to those after pregnancy increased by a factor of 0.0493 ($p < 0.0001$). In 2020, 100% of pregnant people were on ART.

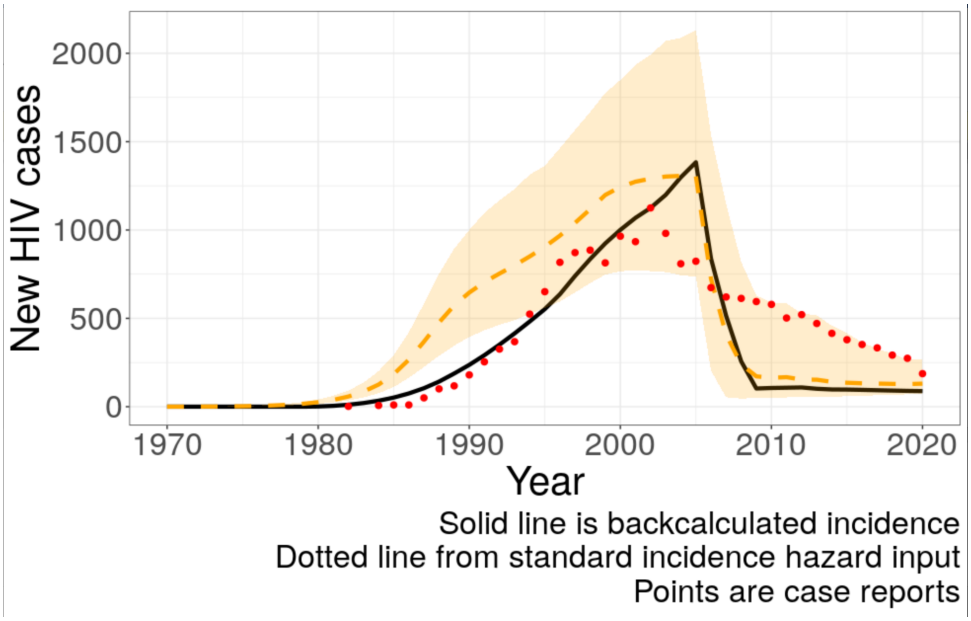

Figure 3a. HIV infections in children under five

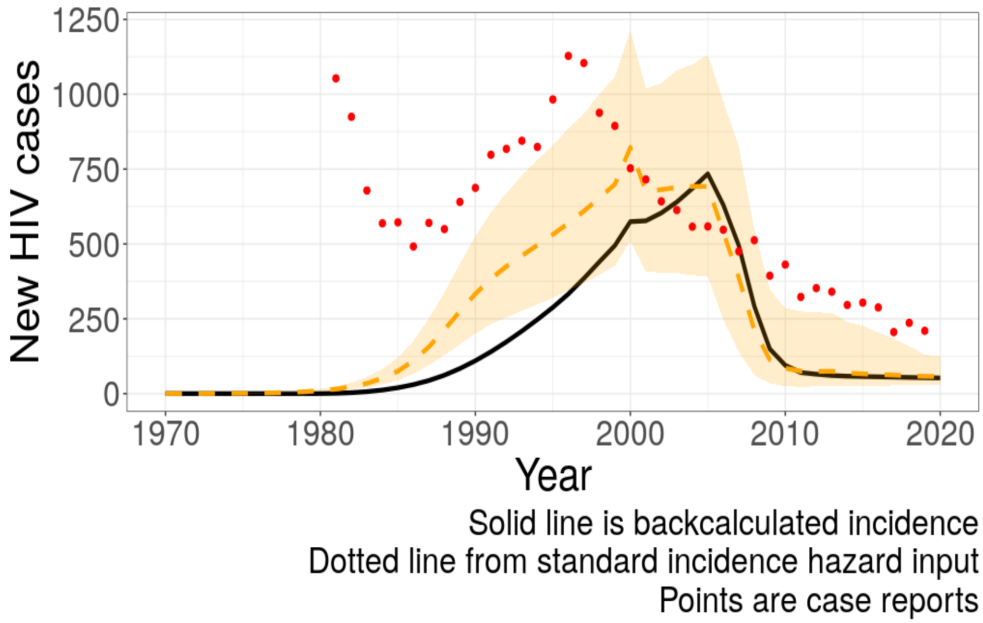

Figure 3b. Deaths due to HIV/AIDS in children under five

Modifying PMTCT and maternal prevalence had a large impact on model fit to both case report data and VR data (Figure 3a). Using the revised model inputs, the peak annual number of new cases among children under 5 was in 2002 (n=1125) and 1996 was the year with the highest number of deaths, with 1128 reported in vital registration data.[7] The peak number of deaths in vital registration data was not reflected in modeled estimates, which estimated that the highest number of deaths in those under five occurred in 2005, with 734 deaths (Figure 3b).

Discussion

In contrast to methods used by UNAIDS and GBD, we used an alternative approach to integrating adult case reports into the Spectrum modeling framework. The method presented here provides insight into how the lag between HIV infection and diagnosis varies across ages, which could have implications for estimating maternal HIV incidence and subsequent pediatric HIV infections over time. Other methods of adjusting case report data to reflect incidence include choosing a standard time lag (such as the five year lag used by the Global Burden of Disease study)[5] or require assumptions surrounding time until case detection and case detection rate provided by countries.[4] Here, we leveraged the benefits of both of these options and improved upon them; regarding the former, this analysis uses an adoptable and adaptable, yet empirically driven, method to inform incidence hazard whereas in terms of the latter we utilize data on CD4 progression to inform case detection.[7,11]

Using CD4 stratified case reports provides insight into how time until diagnosis varies spatially and temporally. Previous work has established that CD4 count are a useful proxy for disease progression,[12] and CD4 at diagnosis has historically been much lower in the Latin America and Caribbean GBD super region than in other super regions, such as Western and Eastern Europe.[4] This suggests that inferring incidence from case reports can not be done blindly; however, even a limited time series of CD4 stratified case reports can be successful in refining adult incidence hazard for Spectrum. Countries

without CD4 stratified case reports could utilize the distribution of lag between infection and diagnosis in adjacent countries to refine estimates.

The incidence inference in our analysis did not result in ubiquitous improvements across disease measures and populations. Our high case detection rate of 93% suggests we may be underestimating the number of undiagnosed infections across ages. The percent of people living with HIV/AIDS who know their status in Latin America is estimated to be much lower than 93%, with many countries in the region having less than 60% of people living with HIV knowing their status in 2018.[4,13] HIV/AIDS mortality in people who do not know their status would be much higher, and could bring our mortality estimates closer to those calculated with VR data. Our estimated incidence also does not reflect the known peaks of HIV incidence in Brazil, which consists of a peak in 1997 and increasing HIV 2001 onwards. These findings may be explained by a high HIV prevalence reported among men who have sex with men and people who inject drugs during these time periods. The highest HIV incidence detected in this analysis in 1997 could be the result of secondary transmission from key populations to the general population.

Our analysis produced incidence estimates that were half as large as those derived from case report data for the under five population post 2005. This may indicate that predicted PMTCT treatment coverage was too high during this period. The methods employed here suggest 100% ART coverage in mothers in 2020, which is be implausibly high.[14] Lower than expected incidence estimates may also point to underestimated transmission among people who seroconvert while breastfeeding. In Brazil, there is no guidance on postpartum HIV testing.[15]

This analysis is limited by the lack of incorporation of social determinant factors that may influence HIV prevalence and the likelihood of being tested, and from this diagnosed with, HIV. Access to HIV testing is driven by a variety of economic and social factors, even in antenatal care.[16,17] While there was an attempt to account for unreported HIV cases, the work presented here is limited in that it does not adjust for receipt of antenatal HIV testing. While leveraging case report data and updated treatment seemed to improve incidence estimates, deaths in children under five were not improved. Estimated deaths are much lower than seen in VR data, but were observed over a different time period. The inconsistency between our estimated deaths and VR data could indicate that pediatric ART coverage or assumed adherence is too high throughout the time series; however, this problem is also seen in the pre-ART era, which could point to unobserved infections throughout the time period. This may point to problems when directly inferring pediatric burden from maternal prevalence. While antenatal care screening coverage is high in Brazil, mothers who are not screened may have higher odds of being HIV positive and untreated. The children of these mothers may then be seen in VR data but not incidence data, leading to an underestimation of burden for of orphans and vulnerable children.[15–17]

The methods applied here can be applied to a variety of countries with high-quality case reports and VR. There are currently 130 locations with both case reports and VR data estimated in the GBD study, allowing for the triangulation and validation of estimates of maternal prevalence. These locations accounted for 8.77% (5.59 – 12.84%) of new pediatric cases of HIV.[5] A parsimonious approach to data collection and modeling allows for more accurate estimates. This gives countries the opportunity to improve their estimates with the data that is currently available rather than having to expand data collection or develop complicated modeling procedures.

Conclusion

More precise HIV prevalence and incidence allow for better prediction of resources needed for HIV care. They also allow for more pointed prevention in subsequent horizontal transmission as those children who perinatally acquire HIV become sexually active. Outside of HIV measurement and estimation, establishing an informed estimate of lag between HIV infection and diagnosis can serve as a proxy for health care utilization. Locations or time periods with longer lags between infection and diagnosis may have healthcare systems that are more difficult to access. This information can also be used to consider which populations are not being tested for HIV but should be more frequently, in terms of age and sex. Closing the gap between observation and estimation also allows for more detection of incident and prevalent cases before they are observed in VR data. This has huge implications at the personal level, where improving incidence and prevalence estimates can help prevent HIV or allow people living with HIV to live longer, healthier lives.

Works Cited

1 AIDS Cases – Since 1980 (SINAN) – DATASUS. https://datasus.saude.gov.br/acesso-a-informacao/casos-de-aids-desde-1980-sinan/ (accessed May 17, 2022).

2 MAHY M, PENAZZATO M, CIARANELLO A, *et al.* Improving estimates of children living with HIV from the Spectrum AIDS Impact Model. *AIDS* 2017; **31**: S13–22.

3 Stover J, Glaubius R, Mofenson L, *et al.* Updates to the Spectrum/AIM model for estimating key HIV indicators at national and subnational levels. *AIDS* 2019; **33 Suppl 3**: S227–34.

4 Mahiane SG, Marsh K, Glaubius R, Eaton JW. Estimating and projecting the number of new HIV diagnoses and incidence in Spectrum's case surveillance and vital registration tool. *AIDS* 2019; **33**: S245.

5 Vos T, Lim SS, Abbafati C, *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 2020; **396**: 1204–22.

6 Govender S, Otwombe K, Essien T, *et al.* CD4 Counts and Viral Loads of Newly Diagnosed HIV-Infected Individuals: Implications for Treatment as Prevention. *PLOS ONE* 2014; **9**: e90754.

7 Kyu HH, Jahagirdar D, Cunningham M, *et al.* Accounting for misclassified and unknown cause of death data in vital registration systems for estimating trends in HIV mortality. *Journal of the International AIDS Society* 2021; **24**: e25791.

8 Brazil Mortality Information System - Deaths 2019. Rio de Janeiro, Brazil: Ministry of Health (Brazil).

9 Glaubius R, Kothegal N, Birhanu S, *et al.* Disease progression and mortality with untreated HIV infection: evidence synthesis of HIV seroconverter cohorts, antiretroviral treatment clinical cohorts and population-based survey data. *Journal of the International AIDS Society* 2021; **24**: e25784.

10 Szwarcwald C, Souza-Júnior P, Pascom A, Ferreira O. Results from a Method for Estimating HIV Incidence Based on the First Cd4 Count among Treatment-Naïve Cases: Brazil, 2004-2013. *Journal of AIDS & Clinical Research* 2016; **7**. DOI:10.4172/2155-6113.1000627.

11 Mahiane SG, Marsh K, Grantham K, Crichlow S, Caceres K, Stover J. Improvements in Spectrum's fit to program data tool. *AIDS* 2017; **31**: S23.

12 Estimation of time of HIV seroconversion using a modified CD4 depletion model | PLOS ONE. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0246135 (accessed May 30, 2022).

13 Mangal TD, Pascom ARP, Vesga JF, Meireles MV, Benzaken AS, Hallett TB. Estimating HIV incidence from surveillance data indicates a second wave of infections in Brazil. *Epidemics* 2019; **27**: 77–85.

14 Domingues RMSM, Saraceni V, Leal M do C. Mother to child transmission of HIV in Brazil: Data from the 'Birth in Brazil study', a national hospital-based study. *PLoS One* 2018; **13**: e0192985.

15 Drake AL, Thomson KA, Quinn C, *et al.* Retest and treat: a review of national HIV retesting guidelines to inform elimination of mother-to-child HIV transmission (EMTCT) efforts. *Journal of the International AIDS Society* 2019; **22**: e25271.

16 Freitas CHS de M, Forte FDS, Galvão MHR, Coelho AA, Roncalli AG, Dias SMF. Inequalities in access to HIV and syphilis tests in prenatal care in Brazil. *Cad Saúde Pública* 2019; **35**. DOI:10.1590/0102-311X00170918.

17 Pascom ARP, Meireles MV, Benzaken AS. Sociodemographic determinants of attrition in the HIV continuum of care in Brazil, in 2016. *Medicine (Baltimore)* 2018; **97**: S69–74.