**WEB APPENDIX WITH SUPPLEMENTARY MATERIALS**

**Text S1. Technical description of the model algorithm**

The mathematical simulation model was constructed using the R package ‘gems’ (<http://cran.r-project.org/web/packages/gems/index.html>). The progression of the disease is represented as a directed acyclic graph of disease states and transitions between them. Figure S1 shows a simplified structure of the model. The complete model consists of 26 states. These states and the possible transitions between them are illustrated in Figure S2. The model allows only forward transitions: a transition from state *i* to state *j* is possible only if *i* < *j*. Each transition corresponds to either a hazard function or a time-to-event function. The hazard functions *h* are of the form *h*(*t*; *bl*,*hist*,***p***), where *t* is the time since the beginning of the state, *bl* the set of baseline characters, *hist* the set of all transition times before reaching that state, and ***p*** a set of parameters. If a transition is defined using a hazard function, the cumulative distribution function of time-to-event is given by*:*

The time to transition *T* is then sampled from such distribution. In the second option (time-to-event function) *T* is defined explicitly as *T*(*bl*,*hist*). Definitions by hazard are used for transitions that represent natural progression (virological and immunological failures, immunosuppression and death). Definitions by time to event are used for transitions that represent fixed events (VL/CD4 tests and switching to second-line ART).

The model simulates each child in the cohort individually. Before starting the simulation, the child is assigned baseline characteristics, such as age and gender. The simulation starts always at the first state (1). Transition times are sampled for all possible transitions that can take place from state 1. All transition times are compared, and the smallest of these times will determine when and to which state the child will progress. If none of the sampled transition times is less than the chosen maximum follow-up time (in our study 5 years), the child will remain in state 1 for the entire duration and the simulation is stopped. If the child progresses to another state, the same procedure is repeated: transition times are sampled for all possible transitions from the new state onwards, and the smallest time will determine the next state the child will enter. This procedure is repeated until the child either enters an absorbing state (i.e. a state from which no onward transitions exist; states 25 or 26 in this model), or no transition takes place until the maximum follow-up time of 5 years is reached. When simulation of one child is done, the transition times are stored and the model moves on to the next child. The output of the model is a matrix of all transition times of all simulated children. This is transformed to outcomes of interest using conventional statistical methods.

**Text S2. Technical description of modelling of mortality**

We modelled mortality as the sum of two components: HIV-unrelated background mortality and HIV-related mortality. We assessed the mortality separately in three groups which were defined according to age at ART initiation: <1 year, 1 to <3 years, and >3 years. We used age- and gender- specific estimates from the ASSA2008 demographic model for Africans in the Western Cape 2007 to calculate the expected background mortality given the time-varying age distribution in the cohort data (<http://aids.actuarialsociety.org.za/ASSA2008-Model-3480.htm>). We then calculated hazard functions for expected background mortality () and all-cause mortality observed in our dataset (). Follow-up of children who reached immunosuppression or immunological failure was censored. The HIV-related mortality is given by:

where the hazard function for HIV-related mortality is . Using this relationship and by assuming piecewise constant rates, we obtained a piecewise constant function for stratified by baseline age group (Table S1). Cox regressions were then performed for the entire dataset to calculate mortality hazard ratios for immunosuppression and immunological failure. These conditions were assumed to be time-dependent covariates. The hazard ratios were estimated using follow-up time on first-line therapy only. The HIV-related mortality rates for children without current immunosuppression or immunological failure are shown in table Table S1 and the corresponding hazard ratios in Table 1. As described in Text S1, our model sampled transition times from hazard functions for mortality, with HIV-related death and HIV-unrelated death modelled as separate states (Figure S2). The transition functions from any state to HIV-unrelated death (state 26) were taken directly from the ASSA2008 model – functions only depended on the current age of the child. Transition functions to HIV-related death (state 25) were modelled using the hazard functions and hazard ratios which were obtained as described above: at each time point, the hazard of HIV-related death is the basic hazard (piecewise constant rate, depending on baseline age and time since ART initiation) multiplied with the corresponding hazard ratio if the child is in a state of immunosuppression or immunological failure at that time.

Because of the considerably higher HIV-related mortality in the early months of ART compared with later months, mortality in the entire cohort is highest during the first few months. Figure S4 illustrates this fact by showing modelled mortality by the different monitoring strategies and an additional scenario where we assumed that no monitoring and second-line ART was available.

**Figure S1. Simplified structure of disease progression in the model.** The ‘Low CD4’ stage includes two sub-stages (immunosuppression and immunological failure) which are defined by CD4 percentages for children aged <5 years (10-15%, <10%) and by absolute CD4 cell counts for children aged ≥5 years (100–200 cells/μl, <100 cells/μl). The dashed thick line indicates transitions associated with routine viral load monitoring, the dashed thin line transitions associated with CD4 monitoring and solid lines represent natural disease progression which is common to all strategies. The colour coding of this figure corresponds to Figure S2, where the structure of the model is represented in detail.

First-line ART

Low CD4

Virological failure

Virological failure

 &

Low CD4

Second line ART

Low CD4

Virological failure

Virological failure

 &

Low CD4

**Figure S2. States of the mathematical model.** The model consists of 26 states that represent different vital, virological and immunological stages of children on 1st- and 2nd-line ART. The following abbreviations and definitions are used to describe the states: VF, virological treatment failure (representing a rebound in viral load to >1000 copies/ml); CIS, concordant immunosuppression; DIS, discordant immunosuppression; CIF, concordant immunological failure; DIF, discordant immunological failure. “Immunosuppression” refers to CD4 percentage <15% (children aged <5 years) or absolute CD4 cell count <200/μl (children aged >5 years); “immunological failure” refers to CD4 percentage below 10% (children aged <5 years) or absolute CD4 cell count below 100/μl (children aged >5 years). “Concordant” represents a condition which is caused by virological failure; “discordant” a condition caused by other causes. The colour coding of this figure corresponds to Figure S1.

**Legend:**

|  |  |
| --- | --- |
| ***v*** | Transition is possible in all monitoring strategies |
| **VL** | Transition is possible only in viral load monitoring strategies |
| **CD4** | Transition is possible only in CD4 monitoring strategy |
| **X** | Transition is never possible |



**Figure S3. Predicted** 5-year mean time spent on failing NNRTI- (left) and PI- (right) based antiretroviral therapy according to different monitoring strategies (Analysis I).
ART: antiretroviral therapy; VL: viral load; NNRTI: non-nucleoside reverse transcriptase inhibitor; PI: protease inhibitor.

**Figure S4.** Predicted overall cumulative mortality of children on ART for all strategies as well as in the hypothetical scenario with only first-line ART (“No monitoring”), based on Analysis I.



**Figure S5.** Predicted Kaplan-Meier probability of death from virological failure for all strategies as well as in the hypothetical scenario with only first-line ART (“No monitoring”). This figure represents mortality in children who may benefit from switching to second-line ART. Time is measured from the first episode of virological failure.

**Table S1. HIV-related mortality rates stratified by age at and time since ART start.** Rates are given per 1000 person-years. The rates are calculated from data of the seven South African cohorts by deducting the expected HIV-unrelated mortality from the observed mortality. The rates are multiplied by the hazard ratios specified in Table 1 for children in states of immunosuppression or immunological failure.

|  |  |
| --- | --- |
| **Months since ART initiation**  | **Age at baseline (years)** |
| **<1** | **1 - <3** | **>3** |
| **0 - <5** | 231.296 | 53.111 | 36.544 |
| **5 - <6** | 27.038 | 53.111 | 36.544 |
| **6 - <10** | 27.038 | 15.512 | 16.281 |
| **10 - <12** | 27.050 | 15.512 | 16.281 |
| **12 - <14** | 27.050 | 12.932 | 9.542 |
| **14 - <18** | 19.768 | 12.932 | 9.542 |
| **18 - <24** | 5.344 | 6.421 | 1.843 |
| **>24** | 5.344 | 4.462 | 1.843 |

**Table S2. Characteristics** **of children included in the parameterization of the model.** All children started antiretroviral therapy in one of seven urban South African sites between 2000 and 2012.

|  |  |
| --- | --- |
| Total number of children | 11903 |
| Follow-up time [years, median (IQR)] | 2.1 (0.7 - 4.2) |
| Gender [count(%)] |   |
|  Male | 5978 (50.2%) |
|  Female | 5925 (49.8%) |
| Age at start of ART [years, median(IQR)] | 3.57 (1 -7.5) |
| Weight-for-age z score\* [median(IQR)] (n = 7287) | -1.17 (-1.95 - -0-32) |
| CD4 cell count\*\* [cells/μl, median (IQR)] (n = 3250) | 231 (80 - 424) |
| CD4 percentage\*\*\* [median (IQR)] (n = 4762) | 15.6 (10 - 22.8) |
| Viral load [ log10 copies/ml, median (IQR) ] (n = 6883) | 5.3 (4.5 - 5.9) |
| First-line regimen [count (%)]  |   |
|  PI-based |   |
|  LPV/r plus at least 2 NRTIs  | 5091 (42.8) |
|  NNRTI-based |   |
|  NVP plus at least 2 NRTIs | 419 (3.5) |
|  EFV plus at least 2 NRTIs | 6393 (53.7) |
| PMTCT exposure [count(%)]  |   |
|  Yes | 1514 (12.7) |
|  No | 2220 (18.7) |
|  Unknown | 8169 (68.6) |
| Confirmed first line virological failures [count (%)](n = 6484)\*\*\*\* |   |
|  Threshold 400 copies/ml | 1805 (27.8) |
|  Threshold 1000 copies/ml | 1507 (23.2) |
| Confirmed first line immunological failures [count](n = 7036)\*\*\*\* | 135 (1.9) |
| Median time between monitoring lab tests [months, median(IQR)] |   |
|  Viral load | 5.7 (4.3-7.2) |
|  CD4 | 6.1 (5.3-7.4) |
| Switches to second line [count (%)] | 842 (7.0) |
|  <5 years\*\* | 320 (4.5) |
|  >5 years\*\*\* | 522 (10.1) |
|   |   |
| ART, antiretroviral therapy; IQR, interquartile range; NNRTI, non-nucleoside reverse transcriptase inhibitor; NRTI, nucleoside reverse transcriptase inhibitor; PI, protease inhibitor; PMTCT, prevention of mother-to-child transmission prophylaxis; LPV/r, ritonavir-boosted lopinavir; NVP, nevirapine; EFV, efavirenz |
| \*Only Children aged <10 years at ART start (n = 10402) |   |
| \*\*Only children aged <5 years at ART start (n=7034) |   |
| \*\*\*Only children aged ≥5 years at ART start (n=4869) |   |
| \*\*\*\*Two consecutive measurements above the threshold within one year. Only patients with at least 48 weeks of follow-up and sufficient number of measurements to detect and confirm failures are included. |

**Table S3.** Outcomes at 5 years in the IeDEA Southern Africa cohort data and the simulation model.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **IeDEA-SA data** | **Predictions for CD4 monitoring** | **Predicitons for viral load monitoring** |
| Median time between measurements [months] | 5.7/6.1\* | **6** | **6** | **12** |
| Mortality [%] | 6.5% | 7.1% | 6.9% | 6.9% |
| Ever experienced immune failures [%] | 1.7% | 1.6% | 1.0% | 1.1% |
| Switched to second line [%] | 11.0% | 1.1% | 13.9% | 12.1% |

Follow-up time in the IeDEA-SA data was censored at 5 years to make estimates comparable with the model outcomes.

\*CD4/viral load measurements