Appendix

Mixed effect models

Mixed effects models are a statistical framework for longitudinal data analysis that provide a picture of the $CD4^+$ T cell response to treatment both in the population as a whole, on average, and in each child individually. This enables quantification of the differences between children attributable to differences in physiology or environment. Mixed effects modelling makes it possible to fit flexible models to longitudinal data with missing values. In addition, the effects of covariates of interest can be investigated. In our mixed-effects model, an individual child's response to time on ART is assumed to follow a trajectory as described by equation (1) in the main text. For each child, the parameters of the model are a combination of a population-average value (the "fixed effect") and a child-specific deviation from this average (the "random effect"):

$$int_i = int_i^f + int_i^r \qquad asy_i = asy_i^f + asy_i^r \qquad c_i = c_i^f e^{c_i^r}$$

where superscripts *f* and *r* represent fixed and random effects, respectively. The fixed effects are modeled as a linear combination of covariate effects:

$$\left(int_i^f, asy_i^f, c_i^f\right) = \boldsymbol{A_i \beta}$$

where $\boldsymbol{\beta}$ is a vector of parameters to be inferred. $\boldsymbol{A_i}$ is a design matrix which links the observations to the fixed effects, as described in the section below (see section titled "covariate model").

The vector of random effects, $\boldsymbol{u_i} = (int_i^r, asy_i^r, c_i^r)$ has a multivariate normal distribution with a mean of 0 and independent of the residual errors (ϵ_{ij}):

$$\boldsymbol{u_i} \sim N(0, \boldsymbol{\Sigma})$$

The residual errors denoted in Equation (1) by ϵ_{ij} are also assumed to be normally distributed with a mean of 0 and variance σ^2:

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

Covariate Model

The final model obtained from the stepwise covariate selection (see Table 2 for more details) can be expressed mathematically as:

$$\begin{pmatrix} asy_i^f \\ int_i^f \\ c_i^f \end{pmatrix} = \begin{pmatrix} asy_{0i} & asy_{A_{gi}} & 0 & 0 \\ int_{0i} & int_{A_{gi}} & 0 & int_{Ei} \\ c_{0i} & 0 & c_{Hi} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ A_{gi} \\ H_i \\ E_i \end{pmatrix}$$

where A_{gi}, H_i and E_i, each represent age at start of ART, HCV status, EPPICC center respectively. The asymptotic model in equation (1) was fitted with random effects added to all three parameters (asy_i, int_i and c_i).

Covariance Matrix

The covariance matrix of the mixed-effects model, which describes the variability within the population in pre-ART and long-term z-scores and their correlation is given below:

	int	*asy*	*c*
int	4.78	0	0
asy	0	1.83	0
c	0	0	0.394

In a separate analysis, the final covariate model described above was modified to include correlation between pre-ART and long-term CD4 z-scores. We found that the inclusion of this correlation made no significant difference to the overall fit of the model.

Age-adjusted CD4 counts

One drawback of the CD4 z-score function is that it does not allow the youngest children to have very low z-scores making it very unstable at low CD4 counts. Hence, low z-scores do not give reliable information for age and based on previous sensitivity analyses, a total of 28 CD4 z-scores below -12 were excluded from our analysis[33].

Supplementary Figure 1

Top: boxplots illustrating the trend in age at start of ART (top left) between mono-infected and co-infected patients. Mono-infected children generally started ART much later than co-infected children. HIV/HCV coinf: HIV/HCV co-infected children, HIV Monoinf: HIV mono-infected children.

Bottom left: diagnostic plot showing good agreement between predicted CD4 z scores and observed CD4 z scores. Bottom right: diagnostic plot inspecting trends in residual error with predicted CD4 z-score. As expected for a good predictive model, no trend is apparent.