

# CD4<sup>+</sup> cell dynamics in untreated HIV-1 infection: overall rates, and effects of age, viral load, sex and calendar time

Anne Cori<sup>a,\*</sup>, Michael Pickles<sup>a,\*</sup>, Ard van Sighem<sup>b</sup>, Luuk Gras<sup>b</sup>, Daniela Bezemer<sup>b</sup>, Peter Reiss<sup>b</sup> and Christophe Fraser<sup>a</sup>

**Background:** CD4<sup>+</sup> cell count is a key measure of HIV disease progression, and the basis of successive international guidelines for treatment initiation. CD4<sup>+</sup> cell dynamics are used in mathematical and econometric models for evaluating public health need and interventions. Here, we estimate rates of CD4<sup>+</sup> decline, stratified by relevant covariates, in a form that is clinically transparent and can be directly used in such models.

**Methods:** We analyse the AIDS Therapy Evaluation in the Netherlands cohort, including individuals with date of seroconversion estimated to be within 1 year and with intensive clinical follow-up prior to treatment initiation. Owing to the fact that CD4<sup>+</sup> cell counts are intrinsically noisy, we separate the analysis into long-term trends of smoothed CD4<sup>+</sup> cell counts and an observation model relating actual CD4<sup>+</sup> measurements to the underlying smoothed counts. We use a monotonic spline smoothing model to describe the decline of smoothed CD4<sup>+</sup> cell counts through categories CD4<sup>+</sup> above 500, 350–500, 200–350 and 200 cells/ $\mu$ l or less. We estimate the proportion of individuals starting in each category after seroconversion and the average time spent in each category. We examine individual-level cofactors which influence these parameters.

**Results:** Among untreated individuals, the time spent in each compartment was 3.32, 2.70, 5.50 and 5.06 years. Only 76% started in the CD4<sup>+</sup> cell count above 500 cells/ $\mu$ l compartment after seroconversion. Set-point viral load (SPVL) was an important factor: individuals with at least 5 log<sub>10</sub> copies/ml took 5.37 years to reach CD4<sup>+</sup> cell count less than 200 cells/ $\mu$ l compared with 15.76 years for SPVL less than 4 log<sub>10</sub> copies/ml.

**Conclusion:** Many individuals already have CD4<sup>+</sup> cell count below 500 cells/ $\mu$ l after seroconversion. SPVL strongly influences the rate of CD4<sup>+</sup> decline. Treatment guidelines should consider measuring SPVL, whereas mathematical models should incorporate SPVL stratification. Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

*AIDS* 2015, **29**:2435–2446

**Keywords:** CD4<sup>+</sup> cell count, HIV, mathematical model, set-point viral load, viral load

## Introduction

Initiation of HAART in HIV-positive individuals is currently recommended to occur once their CD4<sup>+</sup> cell count decreases below 500 cells/ $\mu$ l, according to WHO guidelines [1], although in many settings national

guidelines are still based on the earlier recommendation [2] of CD4<sup>+</sup> 350 cells/ $\mu$ l or less.

Apart from being the main clinical marker for HAART initiation, CD4<sup>+</sup> cell count is also a measure of HIV disease progression. Most mathematical models of HIV

<sup>a</sup>Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK, and <sup>b</sup>Stichting HIV Monitoring, Amsterdam, The Netherlands.

Correspondence to Anne Cori, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London W2 1PG, UK.

E-mail: a.cor@imperial.ac.uk

\* Anne Cori and Michael Pickles contributed equally to the writing of this article.

Received: 15 December 2014; revised: 2 June 2015; accepted: 10 August 2015.

DOI:10.1097/QAD.0000000000000854

ISSN 0269-9370 Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

transmission, including those used in economic evaluation, include a representation of CD4<sup>+</sup> decline over time, usually via compartments: above 500, 350–500, 200–350 and below 200 cells/ $\mu$ l. This structure is based on the historical evolution of the WHO guidelines for treatment initiation.

Models are increasingly playing key roles in policy decision-making in HIV healthcare. Modelling has been influential in the adoption of current WHO treatment guidelines [3] and was used in designing the recent PEPFAR combination prevention trials [4]. Accurately modelling the progression of HIV-infected individuals through these different stages is crucial to provide reliable estimates of incidence, prevalence and number in need of treatment, and hence for robust policy recommendations and assessments.

Although the mean rate of CD4<sup>+</sup> cell count decline has been measured in many studies, the times to reach each of the aforementioned CD4<sup>+</sup> cell count thresholds have, to our knowledge, only been previously analysed in Lodi *et al.* [5]. That study did not explicitly model censoring because of treatment initiation, death or loss to follow-up, and its results had to be processed to be incorporated in mathematical models [6]. Here, we propose to analyse CD4<sup>+</sup> cell count data in a form that can be directly integrated into mathematical and econometric models, and include stratifications that would *a priori* be thought to affect epidemiological dynamics.

Age and sex may influence the CD4<sup>+</sup> cell count at seroconversion [7,8]. Similarly, the rate of decline in CD4<sup>+</sup> cell count depends on a number of factors, including age at seroconversion [7–10] and viral load [10,11]. Set-point viral load (SPVL) is known to also affect the time to AIDS [12–14]. However, to date, no work has quantified how these factors affect the time to reach these CD4<sup>+</sup> thresholds, and therefore mathematical models of HIV transmission seldom stratify disease progression by any of them.

In this study, we use data from a cohort of individuals with date of seroconversion known to be within 1 year to estimate the CD4<sup>+</sup> cell count category at seroconversion and time taken to progress from one CD4<sup>+</sup> category to the next. We explore how these vary according to age at seroconversion, SPVL, sex and HIV transmission route in a treatment-naïve population. These stratifications are to our knowledge novel to this analysis.

## Material and methods

### Data used

We used data from the AIDS Therapy Evaluation in the Netherlands (ATHENA) observational cohort [15]

including HIV-infected patients followed in one of the 27 HIV treatment centres in the Netherlands since 1996. At entry into the cohort, as well as during follow-up visits, clinical, virological and immunological data were collected. The ATHENA cohort includes all patients registered for HIV care in Netherlands, provided they were in clinical care in or after 1996. Data collection was initiated in 1998. All data prior to 1998 were included retrospectively. Data from some patients who died prior to 1996 has also been included retrospectively.

### Data selection

The ATHENA cohort comprises 21 999 individuals in total, of whom we selected the 1039 patients with a seroconversion window (defined by a negative HIV test and a positive HIV test) no longer than a year, at least 6 months between first positive HIV test and antiretroviral therapy (ART) initiation, and at least six CD4<sup>+</sup> cell counts prior to HAART initiation.

Sensitivity analyses were performed for patients who had between three and seven CD4<sup>+</sup> measurements ( $n = 1571$  and  $n = 903$ , respectively); the results as well as further detail on patient selection are shown in the Supplementary Material, <http://links.lww.com/QAD/A766>.

### Ethical approval

Ethical approval in the ATHENA cohort was not obtained as the data used are collected from patients as part of their routine HIV care. Patients can opt out from data collection after being informed by their treating physician of the purpose of collection.

### Progression model

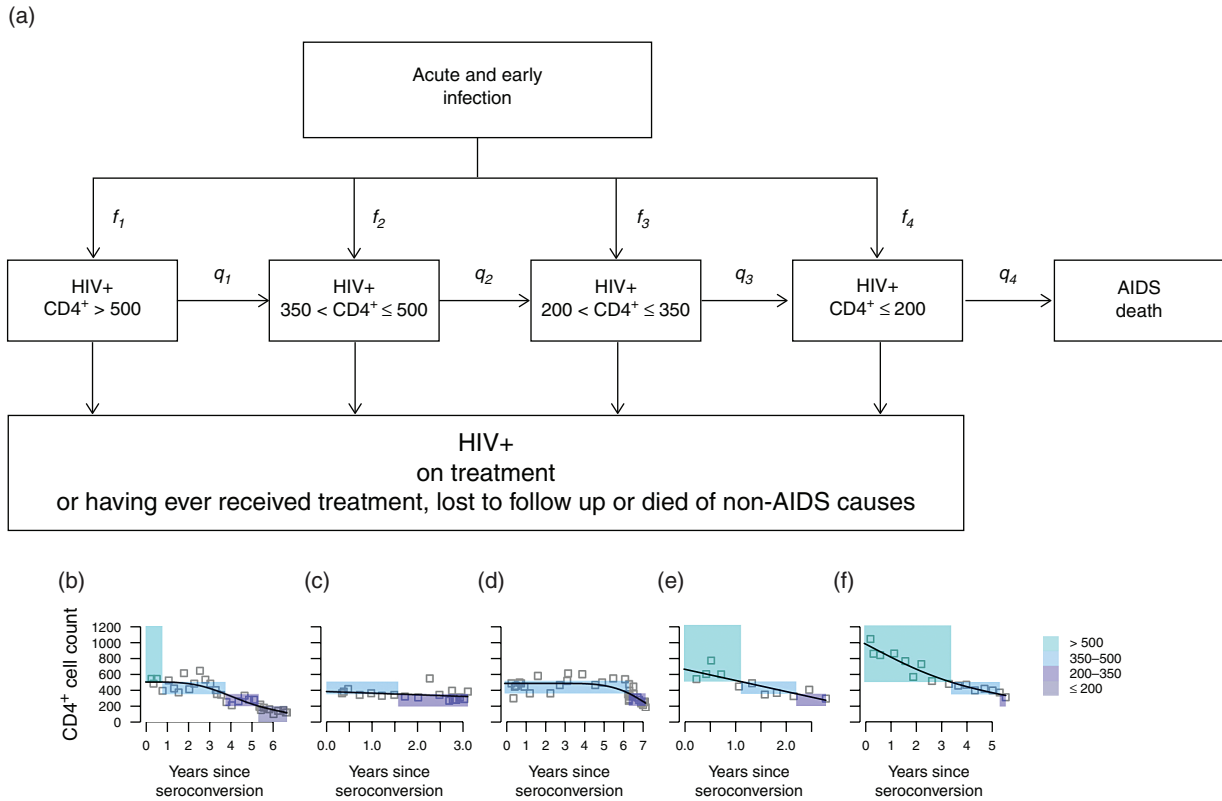
We modelled HIV progression as shown in Fig. 1a. Upon seroconversion, individuals enter one of four categories, defined by their CD4<sup>+</sup> cell count, measured in cells/ $\mu$ l: CD4<sup>+</sup> > 500,  $350 < CD4^+ \leq 500$ ,  $200 < CD4^+ \leq 350$  and CD4<sup>+</sup>  $\leq 200$ . They then successively progress through decreasing CD4<sup>+</sup> categories, at a rate  $q_k$  ( $k=1, \dots, 4$ ) depending on their current CD4<sup>+</sup> category ( $k$ ). At any stage, individuals may be lost to follow-up or die, from HIV-related or unrelated causes. They can also initiate treatment if eligible according to local guidelines.

### Statistical analysis

In the following, the CD4<sup>+</sup> categories referred to are those defined in the previous subsection and illustrated in Fig. 1a.

#### Smoothed and observed CD4<sup>+</sup> cell counts

Owing to the fact that CD4<sup>+</sup> cell counts are highly variable [16], we hypothesize that the model shown in Fig. 1a best describes the underlying dynamics of smoothed CD4<sup>+</sup> cell counts, and that actual CD4<sup>+</sup> cell counts are subject to additional random variability. We thus focus our survival analysis on smoothed CD4<sup>+</sup> cell counts and estimate additional parameters associated with



**Fig. 1. Model of HIV progression during untreated infection (a), illustrated in five patients (b)–(f).** (a) After infection, individuals first go through a short-lived acute and early HIV stage characterized by high viral loads and rapidly changing CD4<sup>+</sup> cell counts. After this, chronic infection sets in, viral loads stabilize around the set-point value and CD4<sup>+</sup> dynamics follow a steady downwards progression. Individuals enter chronic infection in one of four categories defined by their CD4<sup>+</sup> cell count, measured in cells/ $\mu$ l; individuals may thus progress straight from acute infection to a low CD4<sup>+</sup> cell count, with some probability. They then progress through decreasing CD4<sup>+</sup> categories, before initiating treatment or dying.  $f_k$  is the probability of entering the  $k$ th CD4<sup>+</sup> category upon start of chronic infection,  $q_k$  ( $k = 1, \dots, 3$ ) is the rate of progression from the  $k$ th to the  $(k + 1)$ th CD4<sup>+</sup> category,  $q_4$  the death rate for individuals with AIDS (CD4<sup>+</sup> cell count  $\leq 200$ ). Estimates are stratified by age and SPVL. (b)–(f) Fits of monotonic cubic regression splines for five selected individuals. The coloured rectangles show the classification into CD4<sup>+</sup> categories based on the regression. Note that the x-axes show different time scales.

random variation of actual CD4<sup>+</sup> cell counts relative to their underlying trend.

*Estimation of smoothed CD4<sup>+</sup> cell counts*

We used a smoothing model to describe the dynamics of CD4<sup>+</sup> cell count between seroconversion and initiating HAART. More specifically, for each individual, we fitted a monotonic decreasing cubic regression spline to describe the decline of CD4<sup>+</sup> cell count as a function of time since seroconversion [17] (see Supplementary Material for detail, <http://links.lww.com/QAD/A766>).

We used these predicted trajectories of smoothed CD4<sup>+</sup> cell counts to define, for each individual, their CD4<sup>+</sup> category at the estimated date of seroconversion as well as the times of transition to lower CD4<sup>+</sup> categories. The projected smoothed CD4<sup>+</sup> cell counts after acute infection were used to estimate  $f_k$  as the empirical proportion of individuals entering CD4<sup>+</sup> cell count category  $k$  upon seroconversion. We performed

maximum likelihood parametric exponential survival analyses to estimate  $1/q_k$ , that is the average time for all individuals to transition from one CD4<sup>+</sup> category to the next. A progression endpoint was defined if the smoothed CD4<sup>+</sup> cell count crossed a boundary between CD4<sup>+</sup> cell count categories, whereas a censoring event was defined if the patient started treatment, died or reached their final CD4<sup>+</sup> measurement in the database. Confidence intervals (CIs) were obtained for all parameters by bootstrapping patients. Results were visually compared with nonparametric Kaplan–Meier survival estimates.

*Estimation of set-point viral load*

For each patient, we defined SPVL as the geometric mean of all viral load measurements taken during the set-point window (defined as 6–24 months following the first positive HIV test) and prior to HAART initiation. SPVL was available for 873 of the 1039 individuals considered in the main analysis. Viral load units are per millilitre of

peripheral blood plasma (see Supplementary Material for details, <http://links.lww.com/QAD/A766>).

### Subgroup analyses

The analysis was carried out both unstratified and stratified by each of the following characteristics, in order to assess the influence of these factors on individual clinical progression: SPVL, age at seroconversion, sex, calendar time and HIV transmission route. The categories used for  $\log_{10}$  SPVL (<4, 4–4.5, 4.5–5,  $\geq 5$ ) and age at seroconversion (<30, 30–35, 35–40,  $\geq 40$ ) corresponded approximately to the quartiles of SPVL and age at seroconversion.

### Sensitivity analyses

Extensive sensitivity analyses were performed to assess the robustness of the method. Other regression methods were considered to assess the influence of the regression choice on the results. These included an unconstrained cubic spline and a monotonic linear regression, respectively less and more constrained than the monotonic spline, as well as a model with no smoothing. To analyse the robustness of the results to the minimum number of CD4<sup>+</sup> measurements, the analysis with monotonic cubic splines was repeated with a minimum of five to seven CD4<sup>+</sup> measurements, whereas for the linear regression, we further varied the minimum number of CD4<sup>+</sup> measurements from three to seven. To assess the effects of mono and dual ART, we repeated these analyses for individuals using only CD4<sup>+</sup> measurements prior to the initiation of mono or dual therapy, respectively. We also examined the influence of retrospective inclusion of patients diagnosed before 1996 by comparing individuals recruited after 1996 to the whole study population (see Supplementary Material, <http://links.lww.com/QAD/A766> for full details and results of sensitivity analyses).

### Set-point viral load as a continuous variable

Additional analysis of progression to the next CD4<sup>+</sup> category was carried out by applying a Cox proportional hazards model to estimate the progression rates  $q_k$  as linear functions of SPVL.

### Validation

We used as a validation set the 514 patients with seroconversion window between 1 and 2 years, and at least six CD4<sup>+</sup> cell counts prior to HAART initiation. For each of these patients, we predicted, given the first observed CD4<sup>+</sup> category, the times at which the smoothed CD4<sup>+</sup> cell count would transition from one category to the next. Our predictive model accounted for the potential mismatch between the first observed and smoothed CD4<sup>+</sup> categories as well as the smoothed CD4<sup>+</sup> cell count progression over time. The predicted times of transition were then compared with those directly derived by monotonic spline smoothing of the observed CD4<sup>+</sup> cell counts for each patient (see

Supplementary material for detail, <http://links.lww.com/QAD/A766>).

The analysis was carried out using R version 3.1.0. Unconstrained and monotonic splines were created using the *smooth.spline* function in the *stats* package and the *gam* and *mono.con* functions in the *mgcv* package, respectively. Survival analyses were performed using the *survival* package.

## Results

### Study population

The median age at seroconversion of the 1039 individuals included in the main analysis was 35.4 years (interquartile range 28.8–42.2 years); 66 (6.4%) were women, whereas 867 (83.4%) of the study population were MSM. Seventy-four patients had died by the end of the study period, although only two were still HAART naïve. The median initial CD4<sup>+</sup> cell count was 570 (426–720) cells/ $\mu\text{l}$  and a median of 11 CD4<sup>+</sup> measurements were taken over 2.9 years. Only 433 patients had a recorded HIV subtype, and of these, 397 (91.7%) were infected with subtype B.

SPVL was available for 873 (84.0%) patients and based on 1–21 measurements. The median SPVL was 4.6 (4.0–4.9)  $\log_{10}$  copies/ml. Across the cohort, SPVL varied over time (Figure S2, <http://links.lww.com/QAD/A766>) [18]. Mean SPVL in patients who seroconverted during 1995–2000 was 3.9  $\log_{10}$  copies/ml, significantly lower than before 1995 (4.3  $\log_{10}$  copies/ml, *T*-test *P* value <0.01) and after 2000 (4.5  $\log_{10}$  copies/ml, *T*-test *P* value <0.001).

Figure S3, <http://links.lww.com/QAD/A766>, shows the distribution of all individual viral load measurements throughout the set-point window, stratified by SPVL and CD4<sup>+</sup> cell count. Most viral load measurements of a patient lie in the range of their assigned SPVL category, even when the patient progresses to lower CD4<sup>+</sup> categories, showing that SPVL is a good predictor of viral load throughout untreated chronic infection. Only in patients with SPVL below 4  $\log_{10}$  copies/ml, do individuals with low CD4<sup>+</sup> cell counts appear to have on average lower viral load measurements.

### Individual CD4<sup>+</sup> cell dynamics

Figure 1b–f shows, for five selected individuals, their observed and smoothed CD4<sup>+</sup> cell counts (see Supplementary Material for more individuals, <http://links.lww.com/QAD/A766>). Coloured rectangles show the CD4<sup>+</sup> category each individual is inferred to be in over time, from seroconversion onwards, by the smoothed model. The figure illustrates how variable the CD4<sup>+</sup> dynamics are from one individual to the next, both in

**Table 1. Observed versus smoothed CD4<sup>+</sup> category (defined as  $p_{jk}$  in the text).**

	Smoothed CD4 <sup>+</sup> > 500	Smoothed CD4 <sup>+</sup> 350–500	Smoothed CD4 <sup>+</sup> 200–350	Smoothed CD4 <sup>+</sup> ≤ 200
Observed CD4 <sup>+</sup> > 500	4853 (83.98%)	910 (18.69%)	24 (0.85%)	1 (0.16%)
Observed CD4 <sup>+</sup> 350–500	855 (14.79%)	3122 (64.12%)	566 (20.11%)	8 (1.29%)
Observed CD4 <sup>+</sup> 200–350	65 (1.12%)	815 (16.74%)	1961 (69.66%)	66 (10.61%)
Observed CD4 <sup>+</sup> ≤ 200	6 (0.10%)	22 (0.45%)	264 (9.38%)	547 (87.94%)
Total	5779 (100%)	4869 (100%)	2815 (100%)	622 (100%)

Each column in the table shows the distribution of observed CD4<sup>+</sup> cell count for a given category of the smoothed CD4<sup>+</sup>.

terms of the CD4<sup>+</sup> levels immediately after seroconversion and the rate of CD4<sup>+</sup> decline. Compare, for instance, the individual shown in Fig. 1c, who starts with CD4<sup>+</sup> cell count just above 350 cells/μl, with the individual shown in Fig. 1f who still has CD4<sup>+</sup> cell count above 500 cells/μl 3 years after seroconversion.

Table 1 shows the estimates of the distributions  $p_{jk}$ , defined as the proportion of CD4<sup>+</sup> cell counts that are in category  $j$  when the smoothed CD4<sup>+</sup> cell count is in category  $k$ . Smoothed and observed CD4<sup>+</sup> cell counts tend to agree well for the highest and lowest CD4<sup>+</sup> cell count categories ( $p_{11} = 84.0\%$  and  $p_{44} = 87.9\%$ ), but there is considerable misclassification between the second

and third categories ( $p_{23} = 20.1\%$  and  $p_{32} = 16.7\%$ ). In this regard, the recent change in WHO guidelines to start treatment at CD4<sup>+</sup> cell count of 500 cells/μl or less is welcomed, as it avoids arbitrary classification at stage 2 or 3 driven by random variations in CD4<sup>+</sup> cell counts. Our analysis suggests that classifications of CD4<sup>+</sup> cell count of 200 cells/μl or less (roughly corresponding to AIDS) and CD4<sup>+</sup> cell counts above 500 cells/μl are less subject to the vagaries of chance than the classification of CD4<sup>+</sup> stage as above or below 350 cells/μl.

**Population level model of CD4<sup>+</sup> cell dynamics**

Table 2 shows the estimated mean time of progression to the next CD4<sup>+</sup> stage ( $q_k$ ) as well as the estimated

**Table 2. Estimates of mean time  $1/q_i$  (in years) spent in CD4<sup>+</sup> compartment (>500, 350–500, 200–350, ≤ 200 cells/μl), fraction of individuals  $f_i$  starting in each CD4<sup>+</sup> compartment and expected time (in years) to reaching CD4<sup>+</sup> cell count of 200 cells/μl, given current stage of infection.**

Average time spent in each CD4 <sup>+</sup> category		$1/q_1$	$1/q_2$	$1/q_3$	$1/q_4$
Unstratified		3.32 (3.07–3.58)	2.70 (2.48–2.94)	5.50 (4.69–6.54)	5.06 (3.61–7.29)
Stratified by SPVL	$\log_{10}$ SPVL < 4.0	5.35 (4.56–6.37)	3.66 (2.98–4.53)	7.62 (5.04–13.69)	6.59 (3.28–12.87)
	$\log_{10}$ SPVL 4.0–4.5	3.12 (2.68–3.64)	3.09 (2.65–3.64)	8.39 (5.46–15.55)	3.26 (1.43–6.09)
	$\log_{10}$ SPVL 4.5–5.0	2.35 (2.08–2.64)	2.32 (1.98–2.72)	6.57 (4.73–10.22)	9.71 (4.41–23.64)
	$\log_{10}$ SPVL ≥ 5.0	1.51 (1.28–1.76)	1.44 (1.22–1.69)	2.93 (2.12–4.19)	2.14 (1.32–3.59)
Stratified by SC age	SC age < 30	3.83 (3.27–4.44)	3.26 (2.77–3.78)	7.23 (5.31–10.23)	9.28 (4.55–23.90)
	SC age 30–35	3.68 (3.08–4.47)	2.79 (2.28–3.45)	6.05 (4.42–8.94)	9.37 (3.31–20.70)
	SC age 35–40	3.36 (2.89–3.91)	2.76 (2.31–3.32)	4.02 (2.71–6.07)	3.26 (1.92–6.16)
	SC age ≥ 40	2.67 (2.34–3.02)	2.17 (1.89–2.48)	4.75 (3.64–6.62)	2.57 (1.69–4.02)
Proportion in each CD4 <sup>+</sup> category after seroconversion		$f_1$	$f_2$	$f_3$	$f_4$
Unstratified		0.76 (0.73–0.78)	0.19 (0.17–0.22)	0.05 (0.03–0.06)	0.00 (0.00–0.01)
Stratified by SPVL	$\log_{10}$ SPVL < 4.0	0.86 (0.81–0.91)	0.11 (0.07–0.16)	0.02 (0.00–0.04)	0.00 (0.00–0.02)
	$\log_{10}$ SPVL 4.0–4.5	0.78 (0.72–0.83)	0.19 (0.14–0.25)	0.03 (0.01–0.05)	0.00 (0.00–0.00)
	$\log_{10}$ SPVL 4.5–5.0	0.74 (0.69–0.79)	0.21 (0.16–0.26)	0.05 (0.03–0.08)	0.00 (0.00–0.01)
	$\log_{10}$ SPVL ≥ 5.0	0.71 (0.64–0.77)	0.25 (0.19–0.31)	0.04 (0.02–0.07)	0.00 (0.00–0.00)
Stratified by SC age	SC age < 30	0.72 (0.67–0.77)	0.21 (0.17–0.26)	0.06 (0.03–0.09)	0.01 (0.00–0.02)
	SC age 30–35	0.77 (0.71–0.83)	0.17 (0.12–0.23)	0.05 (0.02–0.09)	0.01 (0.00–0.02)
	SC age 35–40	0.77 (0.71–0.83)	0.18 (0.13–0.24)	0.04 (0.02–0.07)	0.00 (0.00–0.00)
	SC age ≥ 40	0.77 (0.73–0.82)	0.20 (0.15–0.24)	0.03 (0.01–0.05)	0.00 (0.00–0.01)
Average time to reaching CD4 <sup>+</sup> 200		From seroconversion	From CD4 <sup>+</sup> > 500	From CD4 <sup>+</sup> 350–500	From CD4 <sup>+</sup> 200–350
Unstratified		10.54 (9.68–11.60)	11.51 (10.59–12.58)	8.19 (7.32–9.26)	5.50 (4.69–6.54)
Stratified by SPVL	$\log_{10}$ SPVL < 4.0	15.76 (12.89–21.96)	16.62 (13.61–22.73)	11.27 (8.53–17.12)	7.62 (5.04–13.69)
	$\log_{10}$ SPVL 4.0–4.5	13.82 (10.75–21.03)	14.60 (11.46–21.76)	11.48 (8.43–18.73)	8.39 (5.46–15.55)
	$\log_{10}$ SPVL 4.5–5.0	10.48 (8.54–14.07)	11.25 (9.28–14.97)	8.90 (6.98–12.48)	6.57 (4.73–10.22)
	$\log_{10}$ SPVL ≥ 5.0	5.37 (4.49–6.63)	5.88 (4.91–7.20)	4.37 (3.49–5.67)	2.93 (2.12–4.19)
Stratified by SC age	SC age < 30	12.95 (10.89–16.21)	14.32 (12.14–17.60)	10.49 (8.42–13.52)	7.23 (5.31–10.23)
	SC age 30–35	11.49 (9.57–14.40)	12.53 (10.48–15.54)	8.84 (7.01–11.58)	6.05 (4.42–8.94)
	SC age 35–40	9.25 (7.76–11.44)	10.13 (8.54–12.44)	6.78 (5.30–8.96)	4.02 (2.71–6.07)
	SC age ≥ 40	8.90 (7.72–10.89)	9.60 (8.37–11.63)	6.92 (5.74–8.87)	4.75 (3.64–6.62)

Mean estimates (95% confidence interval) are shown unstratified and stratified by set-point viral load and seroconversion age. SC, seroconversion.

proportion of individuals entering each of the CD4<sup>+</sup> categories upon entering chronic infection ( $f_k$ ) for the whole population, and then stratified by SPVL and by age at seroconversion. We use these results to derive the estimated average time to reach a CD4<sup>+</sup> cell count of 200 cells/ $\mu$ l, a threshold which is used in many settings to define AIDS, from different time points in the course of infection.

These findings were relatively insensitive to the choice of smoothing method used (see Supplementary Material, <http://links.lww.com/QAD/A766>). They differed from estimates obtained when actual CD4<sup>+</sup> measurements rather than smoothed CD4<sup>+</sup> cell counts were used (i.e. we considered the earliest time when CD4<sup>+</sup> measurements fell below each CD4<sup>+</sup> threshold) which, while giving a comparable time to death, consistently underestimated both the time spent with CD4<sup>+</sup> cell count above 350 cells/ $\mu$ l and the proportion of individuals with CD4<sup>+</sup> above 350 cells/ $\mu$ l after seroconversion, while overestimating the time spent with CD4<sup>+</sup> 350 cells/ $\mu$ l or less.

Patients were only included in the analysis if they had a minimum number (six) of CD4<sup>+</sup> cell counts so that smoothing could be reliably performed; the choice of minimum number of individual CD4<sup>+</sup> measurements used to estimate the smoothing model was also seen not to have an effect on the results (Supplementary Material, <http://links.lww.com/QAD/A766>). Similarly, results were similar when considering only CD4<sup>+</sup> measurements prior to mono or dual ART initiation, or when considering only patients recruited after 1996 (Supplementary Material, <http://links.lww.com/QAD/A766>).

SPVL was found to be strongly associated with CD4<sup>+</sup> progression, with individuals with higher SPVL progressing much faster through all CD4<sup>+</sup> stages than those with lower SPVL (Table 2). Overall, individuals with log<sub>10</sub> SPVL of at least 5.0 progress from seroconversion to CD4<sup>+</sup> cell count of 200 cells/ $\mu$ l approximately twice as fast as those with log<sub>10</sub> SPVL below 4.0 (see Table 2).

In the Cox proportional hazards model analysis, we analysed the data under the assumption that the rate (hazard) of progressing from one CD4<sup>+</sup> category to the next was a linear model of SPVL, an assumption we assessed through a Kaplan–Meier analysis stratified into narrow bands of SPVL (Supplementary Material, <http://links.lww.com/QAD/A766>). Using this model, the relative hazard of progressing to the next CD4<sup>+</sup> category was 2.17 (95% CI 1.89–2.49), 1.88 (95% CI 1.61–2.19), 1.96 (95% CI 1.41–2.73) and 1.63 (95% CI 0.77–3.44) for each log<sub>10</sub> increase in SPVL.

Substantial changes have been seen in population-level SPVL over time (Figure S2, <http://links.lww.com/QAD/A766>), and similarly large differences can be

observed in the unstratified CD4<sup>+</sup> progression rates between different time periods. However, when stratified by SPVL, there was little difference remaining in the rate of CD4<sup>+</sup> progression between 1995 and 2000 compared with before or after (Supplementary Material, <http://links.lww.com/QAD/A766>).

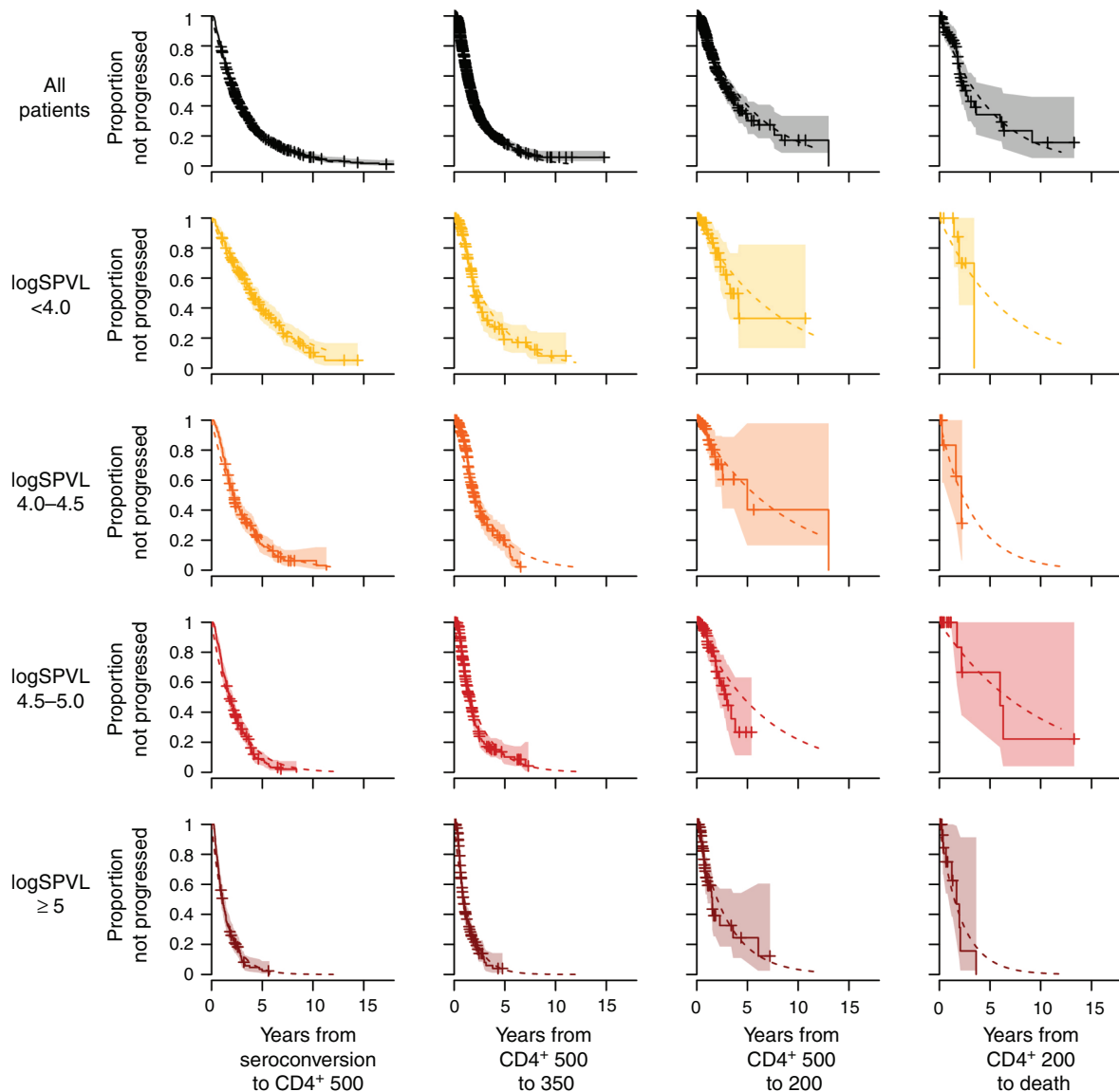
The rate of CD4<sup>+</sup> progression was not found to differ substantially by age at seroconversion, although individuals seroconverting at older ages were found to have a slightly faster progression (Table 2). Comparing MSM/bisexual men to other HIV transmission routes (79% via heterosexual sex, but also including people who inject drugs and one haemophiliac), there was little difference in the rate of CD4<sup>+</sup> progression, although results indicate that MSM/bisexual men may progress more quickly to death once they reach a CD4<sup>+</sup> cell count of 200 cells/ $\mu$ l (Supplementary Material, <http://links.lww.com/QAD/A766>). No significant difference was found in CD4<sup>+</sup> progression between men and women, but there were very few women in the sample and hence only unstratified analysis was possible (Supplementary Material, <http://links.lww.com/QAD/A766>). However, median SPVL was lower in women (4.00 log<sub>10</sub> copies/ml) compared with men (4.57 log<sub>10</sub> copies/ml).

Figure 2 shows, for each CD4<sup>+</sup> category, the survival functions until transition to the next CD4<sup>+</sup> category, for all patients as well as stratified by SPVL.

The nonparametric Kaplan–Meier estimated survival curves are shown alongside the fitted exponential survival curve (i.e.  $e^{-q_k t}$ ,  $k=1..4$ ). This suggests that the exponential survival function provides a very good fit to the nonparametric curve, indicating that the hazard of progression between categories is approximately constant. From this observation, we conclude that the compartmental model shown in Fig. 1a, with constant rates of progression between successive CD4<sup>+</sup> stages, is well supported by the data. In particular, it is well matched to the mean rates of progression, but also provides a good description of interperson variability in progression times.

Figure 3 presents the results of our predictive model, which combines the true CD4<sup>+</sup> progression model and the misclassification model shown in Table 1. The figure shows the predicted future distribution of true CD4<sup>+</sup> categories given current observed CD4<sup>+</sup> category, for all patients and stratified by SPVL. It highlights the large influence of SPVL on patients' future clinical progression.

This predictive model was able to broadly reproduce the true CD4<sup>+</sup> dynamics of the patients selected for validation, given their first observed CD4<sup>+</sup> category (see Figure S6 and Supplementary Material for detail, <http://links.lww.com/QAD/A766>).

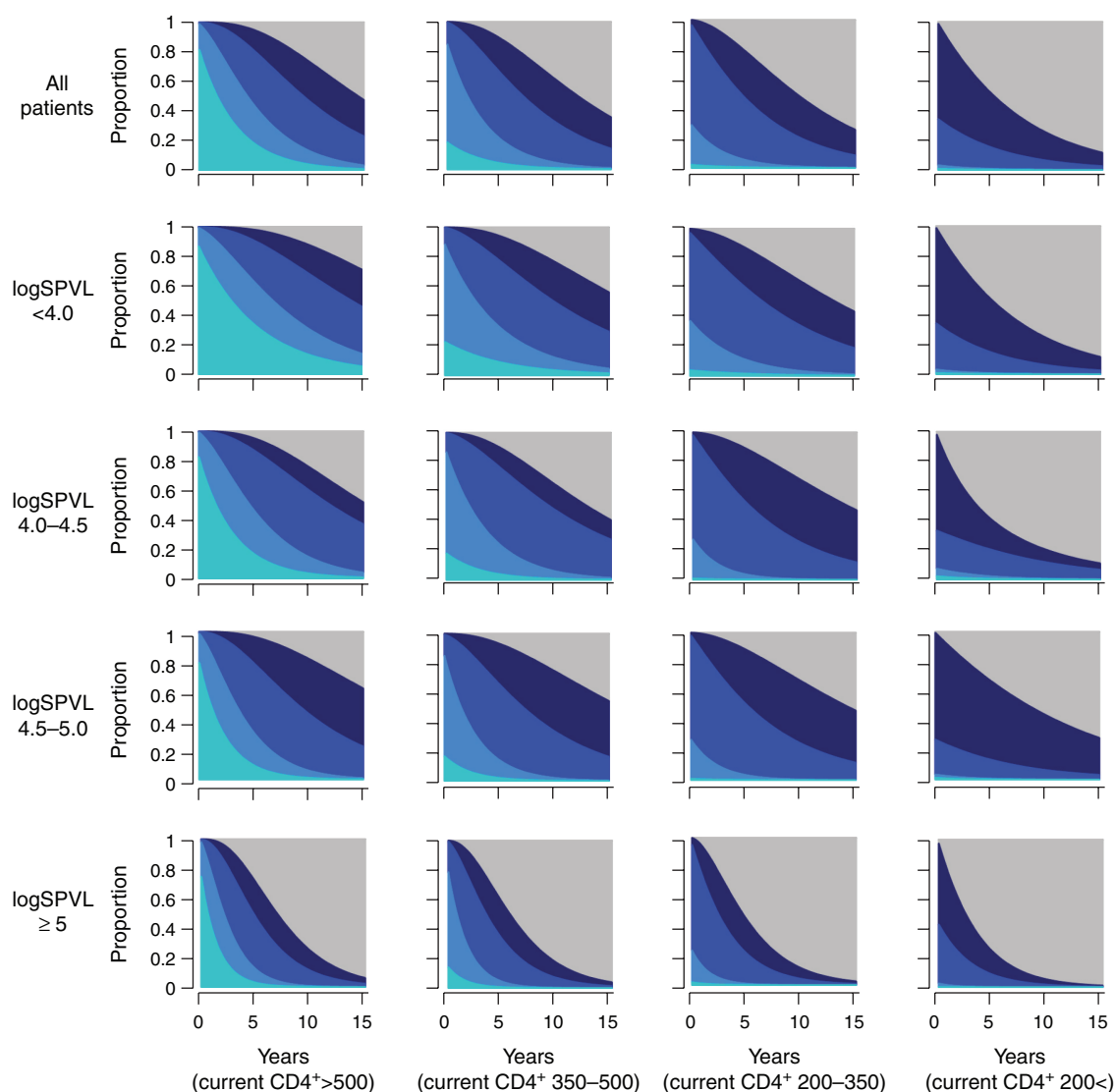


**Fig. 2.** Survival figures, overall and stratified by set-point viral load, showing decline to CD4<sup>+</sup> 500 cells/ $\mu$ l, from 500 to 350, from 350 to 200 and from 200 to death, respectively (darker shades indicate higher SPVL). Crosses and shaded areas indicate nonparametric Kaplan–Meier survival estimates with 95% confidence intervals. Dotted lines show the fitted exponential survival functions.

## Discussion

In this study, we used data from the ATHENA cohort to provide a comprehensive quantitative analysis of CD4<sup>+</sup> cell progression in HIV-infected individuals and the factors affecting the rate of decline. These estimates allow for more accurate long-term projections in mathematical and econometric models. Furthermore, we have found that the structural assumptions in a simple compartmental model of CD4<sup>+</sup> progression (Fig. 1a) provides a very good statistical description of interperson variability in progression rates (Fig. 2) increasing the confidence in such parsimonious modelling approaches.

We showed that a constrained monotonic spline regression model is suitable for smoothing the individual decline of CD4<sup>+</sup> cell counts between the four commonly used CD4<sup>+</sup> categories: CD4<sup>+</sup> >500, 350 < CD4<sup>+</sup> ≤ 500, 200 < CD4<sup>+</sup> ≤ 350 and CD4<sup>+</sup> ≤ 200 cells/ $\mu$ l. We provided estimates of the proportion of individuals who are in each of those categories following seroconversion, as well as estimates of the rates at which individuals progress through the decreasing CD4<sup>+</sup> categories. We found that nearly a quarter of HIV-infected individuals have a CD4<sup>+</sup> cell count below 500 cells/ $\mu$ l immediately following seroconversion. Many mathematical models assume that all individuals start with a CD4<sup>+</sup> cell count above



**Fig. 3. Predicted future distribution of true  $CD4^+$  categories given current observed  $CD4^+$  category (from left to right:  $>500$  cells/ $\mu$ l, 350–500, 200–350 and  $\leq 200$ ) for all patients (top row) and stratified by set-point viral load (second to fifth rows). Shades of blue indicate true  $CD4^+$  category, from light blue for  $CD4^+ > 500$  cells/ $\mu$ l to dark blue for  $CD4^+ \leq 200$  cells/ $\mu$ l; grey indicates death.**

500 cells/ $\mu$ l, and thus they may potentially underrepresent the heterogeneity in  $CD4^+$  progression between individuals, in particular the time to AIDS or death.

In an analysis of data from the CASCADE collaboration, Lodi *et al.* [5] estimated the median time from seroconversion to a  $CD4^+$  cell count below 500, below 350 and below 200 cells/ $\mu$ l as 1.19, 4.19 and 7.93 years, respectively. Gras *et al.* [19] found a similar estimate of over 4 years from seroconversion to a  $CD4^+$  cell count 350 cells/ $\mu$ l for individuals seroconverting before 1996, but a shorter estimate of 2.7 years from seroconversion to  $CD4^+$  cell count 350 cells/ $\mu$ l for those seroconverting after 2002. A number of other studies have reported the median time from estimated date of seroconversion to a  $CD4^+$  cell count below 200 cells/ $\mu$ l. Kilmarx

*et al.* [20] reported 6.9 years among Thai female sex workers, whereas Wandel *et al.* [21] used cohorts from Uganda, Thailand and Côte d'Ivoire and estimated a median of 6.1 years to  $CD4^+$  cell count below 200 cells/ $\mu$ l based on a Weibull survival model, with a shorter duration among the Thai cohorts than the Africans. The difference between our estimates and those of these other studies may be partly explained by their use of the median time. Motivated by the fact that the median time of an exponential process is lower than the mean time, we conducted numerical experiments, and we found a median time of 9.5 years given our model of disease progression.

Cori *et al.* [6] used the data from [5] to parameterize a mathematical model of  $CD4^+$  progression similar to that



presented in this study, finding the total mean time from seroconversion to a CD4<sup>+</sup> cell count of 200 cells/ $\mu$ l was 9.98 years, similar to our estimate of 10.54 years in this analysis. They estimated the proportion of individuals in each CD4<sup>+</sup> category after infection [6] to be  $f_1 = 0.58$ ,  $f_2 = 0.23$ ,  $f_3 = 0.16$  and  $f_4 = 0.03$ . Our analysis led to a slightly higher proportion starting at CD4<sup>+</sup> above 500 cells/ $\mu$ l ( $f_1 = 0.76$ ), and a lower proportion portion starting at  $200 < \text{CD4}^+ \leq 350$  cells/ $\mu$ l ( $f_3 = 0.05$ ) but similar values for the two remaining proportions (Table 2). The average times to progression to the next CD4<sup>+</sup> category estimated in [6] were  $1/q_1 = 6.37$ ,  $1/q_2 = 2.86$ ,  $1/q_3 = 3.54$  and  $1/q_4 = 2.30$  years, respectively. These estimates are quite different from ours ( $1/q_1 = 3.32$ ,  $1/q_2 = 2.70$ ,  $1/q_3 = 5.50$  and  $1/q_4 = 3.54$  years, see Table 2), although in the absence of CIs in [6], it is difficult to assess how significant these differences are. However, Cori *et al.* [6] based their analysis on a 3-year seroconversion window, much longer than the 1-year window used in this study, and this may partly account for the lower fraction of individuals starting at high CD4<sup>+</sup> cell counts found previously.

Importantly, although we have found SPVL to be highly influential in terms of CD4<sup>+</sup> decline, none of these previous studies have reported the population distribution of SPVL, nor have they stratified their analyses by SPVL. Two studies looking at the average rate of CD4<sup>+</sup> decline in South Africa [22] and in Europe [10] similarly found that SPVL was the most influential factor with sex and age playing smaller roles.

It is well known that SPVL and CD4<sup>+</sup> progression are tightly associated [10,11,23]. Here, we provide, to our knowledge, the first analysis which quantifies the association by CD4<sup>+</sup> stage. Our results have multiple direct implications. Clinically, this highlights the value of monitoring viral load. Current WHO guidelines for patient monitoring in pre-ART care do not take SPVL into account. However, patients with higher SPVL prior to initiating treatment will progress faster, and so may need to be monitored more regularly than those with low SPVL (see Fig. 3). For mathematical modelling, it may be important to incorporate SPVL stratification, especially as mean population SPVL has changed over time [18,24] and thus the average rate of progression between CD4<sup>+</sup> stages in the population may have varied over time, consistent with findings from Gras *et al.* [19].

We found some suggestion that individuals seroconverting at older ages may progress more quickly to lower CD4<sup>+</sup> categories, in agreement with the previous studies [7–10], but the effect was small, consistent with [19]. The ATHENA cohort contains mostly MSM, as this is the group with the highest burden of HIV in the Netherlands, and thus we had limited ability to distinguish differences by sex and transmission route. Most patients with recorded subtype were of subtype B, and we were

unable to examine differences between viral subtypes which have been seen in the previous studies in CD4<sup>+</sup> dynamics [25–28] and SPVL [28]. Ethnicity was not routinely recorded in the ATHENA cohort, and thus we could not test differences between ethnic groups [8,9,29,30], and in particular assess the finding by Muller *et al.* that the rate of CD4<sup>+</sup> decline depended strongly on SPVL in patients of European descent, but not in patients of African descent [30]. Extending the study performed here to a wider range of similarly well characterized cohorts will enable a more robust assessment of the extent of variation between populations.

A strength of our study lies in that we have challenged our model through extensive sensitivity analyses. Although all sensitivity analyses gave broadly consistent results, our unstratified estimates were slightly sensitive to the minimal requested number of CD4<sup>+</sup> cell counts, with faster progression inferred when including individuals with few CD4<sup>+</sup> measurements (see Table S4, <http://links.lww.com/QAD/A766>). However, the estimates stratified by SPVL were insensitive to the number of CD4<sup>+</sup> measurements. These results suggest that individuals with very few CD4<sup>+</sup> cell counts may have higher SPVL, leading them to progress faster, hence their small number of CD4<sup>+</sup> measurements. Therefore, although our results stratified by SPVL appear very robust, our unstratified population estimates might be slightly biased towards slow progressors because of our inclusion criteria. Results were very similar whether we considered patients before initiation of HAART, or before initiation of either mono or dual therapy. Similarly, results obtained using the entire dataset and using only individuals diagnosed after 1996 were similar. It is worth noting that our estimates of the average time from a CD4<sup>+</sup> cell count of 200 cells/ $\mu$ l to death are based on very few observations, in particular when stratified by SPVL, as reflected by the wide CIs. Indeed diagnosed HIV-positive individuals with low CD4<sup>+</sup> cell counts are almost always initiated on ART before dying, precluding precise assessment of late-stage progression. Future models using our estimates should account for this uncertainty. In particular, modellers could consider using alternative approaches to modelling the influence of SPVL on late-stage progression, for example, by using the results of our Cox proportional hazard analysis which suggest that, for individuals with CD4<sup>+</sup> less than 200 cells/ $\mu$ l, the death hazard increases by 63% for each log<sub>10</sub> increase in SPVL.

Another novel aspect of our approach is the separation of long-term underlying trends (represented by smoothed CD4<sup>+</sup> cell counts) and random variation (represented by differences between smoothed and observed counts). We hypothesize that, as a result, our findings are more robust to setting and time-specific variation in frequency of CD4<sup>+</sup> testing and can be more readily extrapolated to different settings. Our validation analysis confirmed the validity of our model and our CD4<sup>+</sup> progression

estimates, and indicates that these could be used to predict, at the population level, future true CD4<sup>+</sup> trajectories given current observed CD4<sup>+</sup> cell count distributions.

Our results suggest that SPVL could also be considered in clinical guidelines (see Table 2 and Fig. 3). Current guidelines for pre-ART monitoring recommend a CD4<sup>+</sup> test every 6–12 months, irrespective of SPVL. We estimated that 28% (respectively 38%) of individuals with observed CD4<sup>+</sup> cell count above 500 cells/ $\mu$ l would progress to a true CD4<sup>+</sup> cell count below 500 cells/ $\mu$ l after 6 months (respectively 12 months). However, 43% (respectively 59%) of those with SPVL of at least 5.0 log<sub>10</sub> copies/ml would have already progressed by then. Our results suggest that if the patient's SPVL is known, the frequency of CD4<sup>+</sup> monitoring could be adapted to account for that, with a CD4<sup>+</sup> cell count test every 13–23, 6–12, 3–8 and 2–4 months for individuals with SPVL below 4.0, 4.0–4.5, 4.5–5.0 and at least 5.0 log<sub>10</sub> copies/ml, respectively (see Supplementary Material for detail, <http://links.lww.com/QAD/A766>). These proposed delays are only indicative and should be tailored to each patient in particular given how close their observed CD4<sup>+</sup> cell count is to the threshold 500 cells/ $\mu$ l. One should also keep in mind that 16% of individuals with observed CD4<sup>+</sup> cell count above 500 cells/ $\mu$ l may in fact already have their true CD4<sup>+</sup> cell count below 500 cells/ $\mu$ l, and therefore are only considered ineligible for ART initiation because of stochastic fluctuations in CD4<sup>+</sup> cell counts. Further studies should examine, in light of these results, the costs and benefits of pre-ART monitoring strategies including viral load testing in different settings, to assess the potential cost-effectiveness of such approaches at a population level.

Our analysis has highlighted the fundamental and well known role of viral load in affecting CD4<sup>+</sup> progression dynamics. It has also confirmed that viral load is very stable during the set-point window, suggesting that viral load itself is a good predictor of CD4<sup>+</sup> progression. We have provided estimates of dependence of CD4<sup>+</sup> progression rates on SPVL, both in a subgroup analysis and in a continuous Cox proportional hazards analysis. We suggest that the effect sizes are sufficiently large that models and analyses should be stratified by SPVL wherever possible, if dynamical considerations are of importance. Clinically, these results are relevant for patient monitoring, which should involve both CD4<sup>+</sup> cell count and viral load testing wherever possible.

## Acknowledgements

A.C., M.P. and C.F. thank the NIH for funding through the NIAID cooperative agreement UM1AI068619. C.F. thanks the European Research Council (PBDR-

339251). The authors thank Chris Wymant, Tara Mangal, Jeff Eaton and Tim Hallett for useful discussions.

The ATHENA database is supported by a grant from the Dutch Ministry of Health, Welfare and Sport and was set up and is maintained by Stichting HIV Monitoring. The ATHENA observational cohort has been made possible by the participating patients and through the collaborative efforts of the following physicians working at Netherlands HIV Treatment Centers:

## Clinical centres

'\*' denotes site coordinating physician.

Academic Medical Centre of the University of Amsterdam: HIV treating physicians: J.M. Prins\*, T.W. Kuijpers, H.J. Scherpbier, J.T.M. van der Meer, F.W.M.N. Wit, M.H. Godfried, P. Reiss, T. van der Poll, F.J.B. Nellen, J.M.A. Lange†, S.E. Geerlings, M. van Vugt, D. Pajkrt, J.C. Bos, W.J. Wiersinga, M. van der Valk, A. Goorhuis, J.W. Hovius. HIV nurse consultants: J. van Eden, A. Henderiks, A.M.H. van Hes, M. Mutschelknauss, H.E. Nobel, F.J.J. Pijnappel, A.M. Westerman. HIV clinical virologists/chemists: S. Jurriaans, N.K.T. Back, H.L. Zaaijer, B. Berkhout, M.T.E. Cornelissen, C.J. Schinkel, X.V. Thomas. Admiraal De Ruyter Ziekenhuis, Vlissingen: HIV treating physicians: M. van den Berge, A. Stegeman. HIV nurse consultants: S. Baas. HIV clinical virologists/chemists: L. Sabbe, J. Goudswaard. Catharina Ziekenhuis, Eindhoven: HIV treating physicians: M.J.H. Pronk\*, H.S.M. Ammerlaan. HIV nurse consultants: E.M.H.M. Korsten-Vorstermans, E.S. de Munnik. HIV clinical virologists/chemists: A.R. Jansz, J. Tjhie. Emma Kinderziekenhuis: HIV nurse consultants: A. van der Plas, A.M. Weijnsfeld. Erasmus Medisch Centrum, Rotterdam: HIV treating physicians: M.E. van der Ende\*, T.E.M.S. de Vries-Sluijs, E.C.M. van Gorp, C.A.M. Schurink, J.L. Nouwen, A. Verbon, B.J.A. Rijnders, H.I. Bax, R.J. Hassing, M. van der Feltz. HIV nurse consultants: N. Bassant, J.E.A. van Beek, M. Vriesde, L.M. van Zonneveld. Data collection: A. de Oude-Lubbers, H.J. van den Berg-Cameron, F.B. Bruinsma-Broekman, J. de Groot, M. de Zeeuw-de Man, M.J. Broekhoven-Kruijne. HIV clinical virologists/chemists: M. Schutten, A.D.M.E. Osterhaus, C.A.B. Boucher. Erasmus Medisch Centrum–Sophia, Rotterdam: HIV treating physicians: G.J.A. Driessen, A.M.C. van Rossum. HIV nurse consultants: L.C. van der Knaap, E. Visser. Flevoziekenhuis, Almere: HIV treating physicians: J. Branger\*. HIV nurse consultant and data collection: C.J.H.M. Duijf-van de Ven. HagaZiekenhuis, Den Haag: HIV treating physicians: E.F. Schippers\*, C. van Nieuwkoop, R.W. Brimicombe. HIV nurse consultants: J.M. van Ijperen. Data collection: G. van der Hut. HIV clinical virologist/chemist: P.F.H. Franck. HIV Focus Centrum (DC Klinieken): HIV treating physicians: A. van Eeden\*. HIV nurse consultants: W. Brokking, M. Groot. HIV clinical virologists/chemists: M. Damen, I.S.

Kwa. Isala Klinieken, Zwolle: HIV treating physicians: P.H.P. Groeneveld\*, J.W. Bouwhuis. HIV nurse consultants: J.F. van den Berg, A.G.W. van Hulzen. Data collection: G.L. van der Blik, P.C.J. Bor. HIV clinical virologists/chemists: P. Bloembergen, M.J.H.M. Wolfhagen, G.J.H.M. Ruijs. Kennemer Gasthuis, Haarlem: HIV treating physicians: R.W. ten Kate\*, R. Soetekouw. HIV nurse consultants: N. Hulshoff, L.M.M. van der Pijlt, M. Schoemaker. Data collection: N. Bermon. HIV clinical virologists/chemists: W.A. van der Reijden, R. Jansen. Leids Universitair Medisch Centrum, Leiden: HIV treating physicians: F.P. Kroon\*, S.M. Arend, M.G.J. de Boer, M.P. Bauer, H. Jolink, A.M. Vollaard. HIV nurse consultants: W. Dorama, C. Moons. HIV clinical virologists/chemists: E.C.J. Claas, A.C.M. Kroes. Maastad Ziekenhuis, Rotterdam: HIV treating physicians: J.G. den Hollander\*, K. Pogany. HIV nurse consultants: M. Kastelijns, J.V. Smit, E. Smit. Data collection: M. Bezemer, T. van Niekerk. HIV clinical virologists/chemists: O. Pontesilli.

Maastricht UMC+, Maastricht: HIV treating physicians: S.H. Lowe\*, A. Oude Lashof, D. Posthouwer. HIV nurse consultants: R.P. Ackens, J. Schippers, R. Vergoossen. Data collection: B. Weijenberg Maes. HIV clinical virologists/chemists: P.H.M. Savelkoul, I.H. Loo. MC Zuiderzee, Lelystad: HIV treating physicians: S. Weijer\*, R. El Moussaoui. HIV Nurse Consultant: M. Heitmuller. Data collection: M. Heitmuller. Medisch Centrum Alkmaar: HIV treating physicians: W. Kortmann\*, G. van Twillert\*, J.W.T. Cohen Stuart, B.M.W. Diederer. HIV nurse consultant and data collection: D. Pronk, F.A. van Truijen-Oud. HIV clinical virologists/chemists: W. A. van der Reijden, R. Jansen. Medisch Centrum Haaglanden, Den Haag: HIV treating physicians: E.M.S. Leyten\*, L.B.S. Gelinck. HIV nurse consultants: A. van Hartingsveld, C. Meerkerk, G.S. Wildenbeest. HIV clinical virologists/chemists: J.A.E.M. Mutsaers, C.L. Jansen. Medisch Centrum Leeuwarden, Leeuwarden: HIV treating physicians: M.G.A. van Vonderen\*, D.P.F. van Houte. HIV nurse consultants: K. Dijkstra, S. Faber. HIV clinical virologists/chemists: J. Weel. Medisch Spectrum Twente, Enschede: HIV treating physicians: G.J. Kootstra\*, C.E. Delsing. HIV nurse consultants: M. van der Burg-van de Plas, H. Heins. Data collection: E. Lucas. Onze Lieve Vrouwe Gasthuis, Amsterdam: HIV treating physicians: K. Brinkman\*, P.H.J. Frissen, W.L. Blok, W.E.M. Schouten, G.E.L. van den Berk. HIV nurse consultants: A.S. Bosma, C.J. Brouwer, G.F. Geerders, K. Hoeksema, M.J. Kleene, I.B. van der Meché, A.J.M. Toonen, S. Wijnands. HIV clinical virologists/chemists: M.L. van Ogtrop. Radboud UMC, Nijmegen: HIV treating physicians: P.P. Koopmans, M. Keuter, A.J.A.M. van der Ven, H.J.M. ter Hofstede, A.S.M. Dofferhoff, R. van Crevel. HIV nurse consultants: M. Albers, M.E.W. Bosch, K.J.T. Grintjes-Huisman, B.J. Zomer. HIV clinical virologists/chemists: F.F. Stelma. HIV clinical pharmacology consultant: D. Burger. Rijnstate, Arnhem:

HIV treating physicians: C. Richter\*, J.P. van der Berg, E.H. Gisolf. HIV nurse consultants: G. ter Beest, P.H.M. van Bentum, N. Langebeek. HIV clinical virologists/chemists: R. Tiemessen, C.M.A. Swanink. Sint Elisabeth Hospitaal, Willemstad, Curaçao: HIV treating physicians: C. Winkel, A. Durand, F. Muskiet, R. Voigt. HIV nurse consultants: I. van der Meer. Sint Lucas Andreas Ziekenhuis, Amsterdam: HIV treating physicians: J. Veenstra\*, K.D. Lettinga. HIV nurse consultants: M. Spelbrink, H. Sulman. Data collection: M. Spelbrink, E. Witte. HIV clinical virologists/chemists: M. Damen, P.G.H. Peerbooms. Slotervaartziekenhuis, Amsterdam: HIV treating physicians: J.W. Mulder, S.M.E. Vrouwenraets, F.N. Lauw. HIV nurse consultants: M.C. van Broekhuizen, H. Paap, D.J. Vlasblom. Data collection: E. Oudmajer Sanders. HIV clinical virologists/chemists: P.H.M. Smits, A.W. Rosingh. Stichting Medisch Centrum Jan van Goyen, Amsterdam: HIV treating physicians: D.W.M. Verhagen. HIV nurse consultants: J. Geilings. St Elisabeth Ziekenhuis, Tilburg: HIV treating physicians: M.E.E. van Kasteren\*, A.E. Brouwer. HIV nurse consultants and data collection: B.A.F.M. de Kruijff-van de Wiel, M. Kuipers, R.M.W.J. Santegoets, B. van der Ven. HIV clinical virologists/chemists: J.H. Marcelis, A.G.M. Buiting, P.J. Kabel. Universitair Medisch Centrum Groningen, Groningen: HIV treating physicians: W.F.W. Bierman\*, H.G. Sprenger, E.H. Scholvinck, S. van Assen, K.R. Wilting, Y. Stienstra. HIV nurse consultants: H. de Groot-de Jonge, P.A. van der Meulen, D.A. de Weerd. HIV clinical virologists/chemists: H.G.M. Niesters, A. Riezebos-Brilman. Universitair Medisch Centrum Utrecht, Utrecht: HIV treating physicians: A.I.M. Hoepelman\*, M.M.E. Schneider, T. Mudrikova, P.M. Ellerbroek, J.J. Oosterheert, J.E. Arends, R.E. Barth, M.W.M. Wassenberg. HIV nurse consultants: D.H.M. van Elst-Laurijssen, L.M. Laan, E.E.B. van Oers-Hazelzet, J. Patist, S. Vervoort. Data collection: H.E. Nieuwenhuis, R. Frauenfelder. HIV clinical virologists/chemists: R. Schuurman, F. Verduyn-Lunel, A.M.J. Wensing. VU Medisch Centrum, Amsterdam: HIV treating physicians: E.J.G. Peters\*, M.A. van Agtmael, R.M. Perenboom, M. Bomers, J. de Vocht. HIV nurse consultants: L.J.M. Elsenburg. HIV clinical virologists/chemists: A.M. Pettersson, C.M.J.E. Vandembroucke-Grauls, C.W. Ang. Wilhelmina Kinderziekenhuis, UMCU, Utrecht: HIV treating physicians: S.P.M. Geelen, T.F.W. Wolfs, L.J. Bont. HIV nurse consultants: N. Nauta.

### Coordinating centre

Stichting HIV Monitoring Director: P. Reiss. Data analysis: D.O. Bezemer, L.A.J. Gras, A.I. van Sighem, C. Smit. Data management and quality control: S. Zaheri, M. Hillebrecht, V. Kimmel, Y. Tong. Data monitoring: B. Lascaris, R. van den Boogaard, P. Hoekstra, A. de Lang, M. Berkhout, S. Grivell, A. Jansen. Data collection: L. de Groot, M. van den Akker, D. Bergsma, C. Lodewijk, R. Meijering, B. Peeck, M. Raethke, C. Ree, R. Regtop, Y.

Ruijs, M. Schoorl, E. Tuijn, L. Veenenbergh, T. Woudstra, Y. Bakker, A. de Jong, M. Broekhoven, E. Claessen. Patient registration: B. Tuk.

### Conflicts of interest

There are no conflicts of interest.

### References

- World Health Organization. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach; 2013. [http://apps.who.int/iris/bitstream/10665/85321/1/9789241505727\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/85321/1/9789241505727_eng.pdf?ua=1). [Accessed 1 September 2015]
- World Health Organization. Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach; 2010. [http://apps.who.int/iris/bitstream/10665/44379/1/9789241599764\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/44379/1/9789241599764_eng.pdf) [Accessed 1 September 2015]
- Hallett TB, Menzies NA, Revill P, Keebler D, Bórquez A, McRobie E, Eaton JW. **Using modeling to inform international guidelines for antiretroviral treatment.** *Aids* 2014; **28** (Suppl 1): S1–S4.
- Boily M-C, Mâsse B, Alsallaq R, Padian NS, Eaton JW, Vesga JF, Hallett TB. **HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention.** *PLoS Med* 2012; **9**:e1001250.
- CASCADE Collaboration in EuroCoord. Lodi S, Phillips A, Touloumi G, Geskus R, Meyer L, Thiébaud R, *et al.* **Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds <200, <350, and <500 Cells/mm(3): assessment of need following changes in treatment guidelines.** *Clin Infect Dis* 2011; **53**:817–825.
- Cori A, Ayles H, Beyers N, Schaap A, Floyd S, Sabapathy K, *et al.* **HPTN 071 (PopART): a cluster-randomized trial of the population impact of an HIV combination prevention intervention including universal testing and treatment: mathematical model.** *PLoS One* 2014; **9**:e84511.
- Touloumi G, Hatzakis A, Rosenberg PS, O'Brien TR, Goedert JJ. **Effects of age at seroconversion and baseline HIV RNA level on the loss of CD4+ cells among persons with hemophilia. Multicenter Hemophilia Cohort Study.** *AIDS* 1998; **12**:1691–1697.
- on behalf of CASCADE Collaboration in EuroCoord and ANRS 1220 Primo-CI Study Group. Pantazis N, Morrison C, Amornkul PN, Lewden C, Salata RA, Minga A, *et al.* **Differences in HIV natural history among African and non-African seroconverters in Europe and seroconverters in sub-Saharan Africa.** *PLoS One* 2012; **7**:e32369.
- the Cape Town AIDS Cohort; Swiss HIV Cohort Study. May M, Wood R, Myer L, Taffé P, Rauch A, Battegay M, Egger M. **CD4(+) T cell count decreases by ethnicity among untreated patients with HIV infection in South Africa and Switzerland.** *J Infect Dis* 2009; **200**:1729–1735.
- Natural History Project Working Group for the Collaboration of Observational, H. I. V. E. R. E. i. E. **Factors associated with short-term changes in HIV viral load and CD4(+) cell count in antiretroviral-naïve individuals.** *AIDS* 2014; **28**:1351–1356.
- Langford SE, Ananworanich J, Cooper DA. **Predictors of disease progression in HIV infection: a review.** *AIDS Res Therapy* 2007; **4**:11.
- Mellors JW, Rinaldo CR Jr, Gupta P, White RM, Todd JA, Kingsley LA. **Prognosis in HIV-1 infection predicted by the quantity of virus in plasma.** *Science* 1996; **272**:1167–1170.
- de Wolf F, Spijkerman I, Schellekens PT, Langendam M, Kuiken C, Bakker M, *et al.* **AIDS prognosis based on HIV-1 RNA, CD4+ T-cell count and function: markers with reciprocal predictive value over time after seroconversion.** *AIDS* 1997; **11**:1799–1806.
- Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP. **Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis.** *Proc Natl Acad Sci USA* 2007; **104**:17441–17446.
- van Sighem A, Smit C, Gras L, Holman R, Stolte I, Prins M, de Wolf F. **Monitoring of Human immunodeficiency virus (HIV) infection in the Netherlands 2011.** [http://www.hiv-monitoring.nl/files/6213/2880/0543/Monitoring\\_Report\\_2011.pdf](http://www.hiv-monitoring.nl/files/6213/2880/0543/Monitoring_Report_2011.pdf)
- Raboud JM, Montaner JS, Conway B, Haley L, Sherlock C, O'Shaughnessy MV, Schechter MT. **Variation in plasma RNA levels, CD4 cell counts, and p24 antigen levels in clinically stable men with human immunodeficiency virus infection.** *J Infect Dis* 1996; **174**:191–194.
- Wood S. *Generalized additive models: an introduction with R.* London: Chapman & Hall/CRC; 2006.
- for the ATHENA National Observational Cohort Study. Gras L, Jurriaans S, Bakker M, van Sighem A, Bezemer D, Fraser C, *et al.* **Viral load levels measured at set-point have risen over the last decade of the HIV epidemic in the Netherlands.** *PLoS One* 2009; **4**:e7365.
- for the ATHENA National Observational Cohort. Gras L, Geskus RB, Jurriaans S, Bakker M, van Sighem A, Bezemer D, *et al.* **Has the rate of CD4 cell count decline before initiation of antiretroviral therapy changed over the course of the Dutch HIV epidemic among MSM?** *PLoS One* 2013; **8**:e64437.
- Kilmarx PH, Limpakarnjanarat K, Kaewkungwal J, Srismith R, Saisorn S, Uthairavavit W, *et al.* **Disease progression and survival with human immunodeficiency virus type 1 subtype E infection among female sex workers in Thailand.** *J Infect Dis* 2000; **181**:1598–1606.
- Eligibility for ART in lower income countries (eART-linc) collaboration. Wandel S, Egger M, Ransin R, Nelson KE, Costello C, Lewden C, *et al.* **Duration from seroconversion to eligibility for antiretroviral therapy and from ART eligibility to death in adult HIV-infected patients from low and middle-income countries: collaborative analysis of prospective studies.** *Sexually Transmitted Infect* 2008; **84** (Suppl 1):i31–i36.
- Martinson NA, Gupte N, Msandiwa R, Moulton LH, Barnes GI, Ram M, *et al.* **CD4 and viral load dynamics in antiretroviral-naïve HIV-infected adults from Soweto, South Africa: a prospective cohort.** *PLoS One* 2014; **9**:e96369.
- Geskus RB, Prins M, Hubert JB, Miedema F, Berkhout B, Rouzioux C, *et al.* **The HIV RNA setpoint theory revisited.** *Retrovirology* 2007; **4**:65.
- Herbeck JT, Müller V, Maust BS, Ledergerber B, Torti C, Di Giambenedetto S, *et al.* **Is the virulence of HIV changing? A meta-analysis of trends in prognostic markers of HIV disease progression and transmission.** *AIDS* 2012; **26**:193–205.
- Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F, *et al.* **Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection.** *J Infect Dis* 2008; **197**:707–713.
- Keller M, Lu Y, Lalonde RG, Klein MB. **Impact of HIV-1 viral subtype on CD4+ T-cell decline and clinical outcomes in antiretroviral naïve patients receiving universal healthcare.** *AIDS* 2009; **23**:731–737.
- for the IAVI Africa HIV Prevention Partnership. Amornkul PN, Karita E, Kamali A, Rida WN, Sanders EJ, Lakhi S, *et al.* **Disease progression by infecting HIV-1 subtype in a seroconverter cohort in sub-Saharan Africa.** *AIDS* 2013; **27**:2775–2786.
- for the CASCADE Collaboration in EuroCoord. Touloumi G, Pantazis N, Pillay D, Paraskevis D, Chaix ML, Bucher HC, *et al.* **Impact of HIV-1 subtype on CD4 count at HIV seroconversion, rate of decline, and viral load set point in European seroconverter cohorts.** *Clin Infect Dis* 2013; **56**:888–897.
- Easterbrook PJ, Farzadegan H, Hoover DR, Palenicek J, Chmiel JS, Kaslow RA, Saah AJ. **Racial differences in rate of CD4 decline in HIV-1-infected homosexual men.** *AIDS* 1996; **10**:1147–1155.
- for the Swiss HIV Cohort Study. Müller V, von Wyl V, Yerly S, Böni J, Klimkait T, Bürgisser P, *et al.* **African descent is associated with slower CD4 cell count decline in treatment-naïve patients of the Swiss HIV Cohort Study.** *AIDS* 2009; **23**:1269–1276.