

Long-term follow-up of HIV-2 related AIDS and mortality in Guinea-Bissau: a prospective open cohort study

Authors

Joakim Esbjörnsson*†, Fredrik Månsson†, Anders Kvist, Zacarias J. da Silva, Sören Andersson, Eva Maria Fenyö, Per-Erik Isberg, Antonio J. Biague, Jacob Lindman, Angelica A Palm, Sarah L. Rowland-Jones, Marianne Jansson, Patrik Medstrand, and Hans Norrgren for the SWEGUB CORE GROUP‡

Affiliations

Department of Laboratory Medicine, Lund University, Lund, Sweden

(Joakim Esbjörnsson, PhD; Eva Maria Fenyö, MD; Marianne Jansson, PhD)

Nuffield Department Medicine, University of Oxford, Oxford, United Kingdom

(Joakim Esbjörnsson, PhD; Sarah L. Rowland-Jones, MD)

Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden

(Fredrik Månsson, MD)

Department of Clinical Sciences Lund, Lund University, Lund, Sweden.

(Anders Kvist, PhD; Jacob Lindman, MD; Hans Norrgren, MD)

National Public Health Laboratory, Bissau, Guinea-Bissau

(Zacarias J. da Silva, PhD; Antonio J. Biague, MD)

Department of Laboratory Medicine, Örebro University, Örebro, Sweden

(Sören Andersson, MD)

Department of Statistics, Lund University School of Economics and Management, Lund, Sweden

(Per-Erik Isberg)

Department of Translational Medicine, Lund University, Lund, Sweden

(Angelica A Palm, PhD; Patrik Medstrand, PhD)

*Corresponding author:

Joakim Esbjörnsson

Department of Laboratory Medicine, Lund University

BMC B13

22184 Lund

Sweden

†These authors contributed equally to this work

‡The Sweden and Guinea-Bissau Collaboration Research Group. All collaborators are listed in the paragraph Members of the SWEGUB CORE group.

Full professors

Sören Andersson

Eva Maria Fenyö (emeritus)

Sarah L. Rowland-Jones

Patrik Medstrand

SUMMARY

Background

Compared with HIV-1, HIV-2 is considered more benign and with less pathogenic consequences for the majority of infected individuals. The aim of this study was to compare the time to AIDS and mortality, and the CD4+T cell dynamics between HIV-1 and HIV-2 single-infected individuals in an open prospective cohort of police officers followed in Guinea-Bissau from 6 February 1990 to 28 September 2013.

Methods

We analyzed longitudinal data from 408 HIV-1 infected individuals (183 infected before and 225 infected after enrollment) and 464 HIV-2 infected individuals (377 and 87) in a cohort with up to 23 years follow-up duration, according to time to AIDS, time to death, and T-cell dynamics. Time of HIV infection was estimated as the mid-time point between last HIV-seronegative and first HIV-seropositive sample. Data from an additional 2,984 HIV uninfected individuals from the same population were analysed to assess the impact of natural mortality on HIV-related mortality.

Findings

The median time from HIV infection to development of AIDS was 6.2 years (95% confidence interval [CI] 5.4-7.1) for HIV-1 infection, and 14.3 years (CI 10.7-18.0) for HIV-2 infection ($p < 0.0001$). The median survival time after HIV infection was 8.2 years (95% CI 7.5-8.9) for HIV-1 infection, and 15.6 years (CI 12.0-19.2) for HIV-2 infection ($p < 0.0001$). Individuals that were HIV-1 or HIV-2 infected before enrollment showed similar results. Comparison with uninfected individuals indicated limited confounding contribution from natural mortality. CD4+ T-cell levels were 1.3 times higher in early infection and declined at a rate 2.3 times

slower in HIV-2 compared with HIV-1 infection. HIV-2 infected individuals developed clinical AIDS at higher CD4+ T-cell levels than HIV-1 infected individuals.

Interpretation

Our results show that both HIV-1 and HIV-2 infected individuals have a high probability of developing and dying from AIDS without antiretroviral treatment.

Funding

Swedish International Development Agency, Swedish Research Council, Swedish Society of Medical Research, Medical Faculty at Lund University, and Region Skåne FoU.

RESEARCH INTO CONTEXT

Evidence before this study

We searched PubMed using the MeSH and search terms “HIV-2”, “disease progression”, “mortality” in appropriate variations and combinations with no language restrictions for studies from Jan 1, 1986 to Mar 1, 2017. The literature review showed that the common consensus is that although HIV-2 infection can lead to HIV-related diseases, the majority of HIV-2 infected individuals do not progress to immunodeficiency and AIDS during follow-up. Moreover, it has been suggested that HIV-2 infection is compatible with a normal lifespan in most infected individuals and that previous mortality rate estimates may have been influenced by natural mortality. Importantly, the literature review also showed that only one previous report followed their study participants from an estimated infection date (33 individuals, followed for a median time of three years). Information on date of infection is particularly important for precise estimates of disease duration and mortality rates in progressive chronic diseases like HIV infection (where time from infection to disease can vary from years to decades between individuals).

Added value of this study

This is the first study to present reliable estimates from time of HIV infection to AIDS or HIV-related death for HIV-2 with comparative estimates for both HIV-1 and HIV-negative individuals from the same population. We also provide novel understanding of how the T-cell dynamics differ between HIV-1 and HIV-2 infection, pointing towards determinants in early infection as having the strongest impact on disease progression.

Implications of all the available evidence

We show that HIV-2 in general is more pathogenic than previously thought and that the majority of HIV-2 infected individuals will progress to HIV-related disease and death without

treatment. Our findings suggest that early treatment initiation should be considered for all HIV-infected individuals, not only those infected by HIV-1. This also reiterates previous calls for the urgency in generating clinical trial data to establish the evidence base for optimal usage of ART in HIV-2 infection.

INTRODUCTION

Two types of human immunodeficiency viruses (HIV) have been discovered since the description of acquired immunodeficiency syndrome (AIDS). Although HIV type 1 (HIV-1) is responsible for most of the HIV pandemic, approximately 1-2 million people are infected with HIV-2 worldwide.¹ HIV-2 was first isolated in individuals from West Africa in 1986, and has since then been found in other parts of the world. In West Africa, previous studies have shown a crossing pattern in prevalence between HIV-1 and HIV-2, where HIV-1 has been the dominating HIV type since around the year 2000.^{2,3} Both viruses share transmission routes, cellular targets, and spectrum of opportunistic infections.^{4,5} HIV-2 infection is, however, characterized by lower transmission rates, longer asymptomatic stages, slower declines in CD4+ T-cell counts, and lower mortality rates.⁶⁻¹³ Without antiretroviral therapy (ART), the vast majority of individuals develop AIDS and die within 3-13 years after HIV-1 infection.¹⁴ Corresponding estimates among HIV-2 infected individuals has not been presented, mainly due to lack of data with both estimated date of HIV-2 infection and long follow-up. Other limitations include few study participants, infrequent follow-up, high enrollment age, or low CD4+ T-cell counts and late-stage disease at enrollment.^{3,15,16} In addition, it has been suggested that only a minority of HIV-2 infected individuals develop AIDS during follow-up, that HIV-2 infection is compatible with a normal lifespan in most infected individuals, and that previous mortality rate estimates are heavily influenced by natural mortality.^{2,3,9,16} Moreover, no large observational and randomized HIV-2 treatment study has been performed and the current guidelines for antiretroviral treatment of HIV-2-positive individuals from the British HIV Association indicate that initiation of antiretroviral therapy should be based on CD4 cell count and clinical status.^{1,15} Hence, comprehensive data on HIV-2 natural disease progression is needed.

METHODS

Study design and participants

The study was approved by the ethical committees of the Ministry of Health in Guinea-Bissau and at Lund University and Karolinska Institute, Sweden. All participants received detailed information about the study before inclusion. All participants that tested HIV positive were counselled and informed about antiretroviral treatment (ART), the national treatment guidelines (following WHO recommendations), and offered ART accordingly. Up to 2011, only oral informed consent was required. Since then, written informed consent was required from all study participants.

An open prospective cohort of police officers in Guinea-Bissau was established the 6 February, 1990. Inclusion of participants continued until 28 September, 2009 and all individuals were followed until 25 February, 2011. However, follow-up of the HIV-1 and HIV-2 positive individuals continued until 28 September, 2013. All persons with a regular employment in the Guinea-Bissau police force were eligible for inclusion (>98% of police officers accepted to be included in the study).

Procedures

Blood samples were collected at enrollment and at scheduled annual follow-up visits at police stations in both urban and rural areas of Guinea-Bissau. HIV testing was performed at the National Public Health Laboratory in Bissau. Body weight was measured and clinical symptoms were classified according to the World Health Organization (WHO) clinical staging system.¹⁷ Mortality reports from the Ministry of Interior were cross-checked by health post staff at the main police station in the capital Bissau. Participants with reported mortality were clinically classified according to reported symptoms before death. ART was introduced in Guinea-Bissau in 2005 through a national treatment programme. The police cohort was

included in the programme in early 2006. Individuals receiving ART were censored from the analysis from the time-point of ART initiation (in total 48 participants of whom 12 became HIV-1 infected after enrollment, 17 were HIV-1 infected at enrollment, six became HIV-2 infected after enrollment, and 13 were HIV-2 infected at enrollment). Date of HIV seroconversion was estimated as the mid-time point between the last HIV-seronegative and the first HIV-seropositive sample.^{18,19} Detailed information of the cohort are available in the Appendix.

Statistical analyses

We analysed longitudinal data from HIV-1 infected individuals and HIV-2 infected individuals according to time to AIDS, time to death, and T-cell dynamics.

Both absolute CD4+ T-cell counts and percentage (CD4%) are reliable immunologic markers of HIV disease progression. In resource-limited settings, the percentage levels can be more suitable due to lower variability and sensitivity to specimen handling, patient age or time of sampling.^{19,20} Accordingly, we present on CD4% and CD8% levels (absolute T-cell counts can be found in Appendix, p. 20-24). Longitudinal T-cell data were analysed by full-factorial linear mixed models with time since infection as covariate.

Kaplan-Meier analyses were performed for progression-time to AIDS and mortality. AIDS was defined according to WHO staging as CD4+ T-cell counts ≤ 200 cells per μL or CD4% counts $\leq 14\%$ (immunological criteria), or AIDS-defining symptoms (WHO 4 clinical criteria, Appendix, p. 5).^{17,19} The most common clinical condition was wasting syndrome. Cases that did not reach AIDS or death during follow-up were right censored at their last clinical examination date. Statistical comparisons were performed by the Log-Rank test. Finally, a

Cox proportional-hazards model was applied, adjusting for sex and age at estimated date of infection or enrollment date for participants HIV-infected before enrollment (age was modelled as a continuous variable). The most complex model with all possible interaction terms was initially applied followed by a step-wise removal of non-significant interaction terms. Because of imbalance in sex representation (Table 1), analyses were also stratified by sex. To control that the proportional-hazards assumption was satisfied, we analysed log-minus-log plots to rule out any crossing of the curves. Statistical comparisons were performed by the Wald test.

Expected survival analysis was performed to disentangle the impact of natural mortality on HIV disease progression. Expected survival, matched by age and sex, was predicted from a Cox proportional-hazards model of the HIV-negative individuals in the cohort using the Hakulinen method.²¹ Weibull distributions were fitted to the survival data to assess differences in shape and scale parameter between HIV-1 and HIV-2 infections. Proportional hazards and accelerated failure time assumptions were checked by parallel straight lines of log-minus-log plots to confirm the appropriateness of the Weibull distributions.

Statistical analyses were performed in IBM® SPSS® Statistics Version 23 (IBM Corporation), and in R 3.2.2 using the “survival”, “flexsurv”, “relsurv” and “yarr” packages.

Role of the funding source

The study was supported by the Department for Research Cooperation (SAREC), Swedish International Development Agency (SIDA), the Swedish Research Council, Swedish Society of Medical Research, Medical Faculty at Lund University (ALF), and Region Skåne FoU. The funders of the study had no role in data collection, data analysis, data interpretation, or writing

of the report. All authors had full access to all the data in the study and the corresponding author had final responsibility for the decision to submit for publication.

RESULTS

Overall, 872 participants were tested HIV-positive during the study period. Of these, 74 participants eligible for ART (8%) remained untreated during follow-up: 19 declined to start treatment even after detailed information about the risks not taking ART; two participants started ART but chose to stop treatment only few weeks after treatment initiation and were therefore considered untreated; 48 participants reached the eligibility criteria at their last visit close to study closure and had therefore not initiated ART at that time; and for five participants information on whether they were offered ART or not was missing.

Two-hundred-and-twenty-five individuals were infected with HIV-1 and 87 individuals with HIV-2 after enrollment over a period of 23 years follow-up in a prospective cohort of police officers in Guinea-Bissau. Age and sex distributions were similar between the HIV-infected groups (Table 1). No evidence (e.g. follow-up duration, follow-up rate, loss-to-follow-up, censoring, or participant characteristics affecting the below analyses were found (Appendix).

The median survival time was 8.2 years (95% confidence interval [CI] 7.5-8.9) in HIV-1 and 15.6 years (CI 12.0-19.2) in HIV-2 infected individuals ($p < 0.0001$, Log-Rank test, Figure 1A). The Hazard Ratio (HR) for mortality among HIV-1 compared with HIV-2 infected individuals was 3.50 in a Cox proportional-hazards model controlling for age and sex (CI 2.22-5.52, $p < 0.0001$, Wald test). Age (HR 1.03 per year of age increase, CI 1.01-1.05, $p = 0.0033$, Wald test) and sex (HR 2.11 for men vs. women, CI 1.18-3.79, $p = 0.012$, Wald test) were independently associated with mortality. Survival analysis based on HIV-related death instead of all-cause mortality gave similar results (Appendix, p. 16).

One-hundred and twenty-one HIV-1 (54%) and 37 HIV-2 infected individuals (43%) developed AIDS during follow-up. The median time to AIDS was 6.2 years (CI 5.4-7.1) in

HIV-1 and 14.3 years (CI 10.7-18.0) in HIV-2 infected individuals ($p < 0.0001$, Log-Rank test, Figure 1B). The HR for time to AIDS among HIV-1 compared with HIV-2 infected individuals was 2.84 in a Cox proportional-hazards model controlling for age and sex (CI 1.91-4.22, $p < 0.0001$, Wald test). Similar to the mortality analysis, age (HR 1.03 per year of age increase, CI 1.01-1.04, $p = 0.0025$, Wald test) and sex (HR 1.33 for men vs. women, CI 1.05-1.71, $p = 0.020$, Wald test) were independently associated with time to AIDS. Separate analysis of HIV-1 and HIV-2 showed significantly higher risk of AIDS among HIV-2 infected men compared with women (HR 5.10, CI 1.22-21.74, $p = 0.027$, corrected for age, Wald test). For HIV-1 infected men and women the difference was less pronounced (HR 1.36, CI 0.81-2.31, $p = 0.25$, Wald test). Log-minus-log plots were analysed for all Cox regressions and consistently showed parallel curves, indicating that the proportional-hazards assumption was satisfied.

During the study period, 183 and 377 individuals were recorded with HIV-1 or HIV-2 infection already at enrollment. These individuals differed in disease stage at enrollment and lacked an estimated HIV-infection date. Nevertheless, HIV-2 infected individuals had longer survival times (15.7 years [CI 13.5-17.9] vs. 6.6 years [CI 5.8-7.4], $p < 0.0001$, Log-Rank test) and times to AIDS (11.2 years [CI 9.9-14.1] vs. 5.2 years [CI 3.9-6.5], $p < 0.0001$, Log-Rank test) from enrollment than HIV-1 infected individuals also in this population (Appendix, p. 18-19).

To assess the impact of high enrollment age and natural mortality³, we compared mortality of HIV-infected individuals with estimated date of infection and HIV-negative individuals in our cohort (Table 1). Kaplan-Meier analysis showed a median mortality age of 51.0 years (CI 48.8-53.2) in HIV-1 infected individuals, 62.9 years (CI 54.3-71.6) in HIV-2 infected

individuals, and 73.7 years (CI 67.7-79.7) in HIV-negative individuals ($p < 0.0001$ for all pairwise comparisons, Log-Rank test). There were significant differences in sex and age distributions between HIV-positive and HIV-negative individuals (Table 1). To adjust for this, we compared observed survival of HIV-infected individuals with expected survival of HIV-negative individuals matched for age and sex (Figure 2A). Graphical comparison shows substantial differences in mortality rates between groups, suggesting marginal impact of natural mortality in the HIV-infected populations throughout the disease course.

Next, we wanted to address the hypothesis that most HIV-2 infected individuals never develop AIDS. The proportion of non- or very slow progressors can be estimated from the tail of the survival curve where the incidence of AIDS and mortality rate drops to the level of uninfected individuals. Despite long follow-up, such a decline in mortality rate is not seen in our cohort (Figure 2A). However, valuable insight can be gained by comparing the survival of HIV-2 with that of HIV-1 infected individuals – for which we know the proportion of non- or slow progressors is very low.^{3,22} Parametric Weibull models of the survival curves for HIV-1 and HIV-2 infected individuals mainly differed in the scale parameter (10.56 [CI 9.68-11.52] for HIV-1 and 19.31 [CI 14.01-26.61] for HIV-2), while the shape parameters were almost identical (HIV-1: 2.21 [CI 1.93-2.53], HIV-2: 2.12 [CI 1.50-2.77], Figure 2B). A model with a common shape parameter but separate time scales provided a better fit (Aikike Information Criterion, AIC=878.6) than a model with separate shape and a common time scale (AIC=916.4). Despite a similar fit, this model should also be preferred over a model with separate shape and separate scale (AIC=880.3) since it is more parsimonious (according to the law of parsimony, Occam's razor). The estimated shape parameters close to two indicate an almost linear increase in risk of death with time from HIV-infection for both groups. Consequently, the density plots peaked marginally before the median survival time with a

slightly right-tailed distribution (Figure 2C). The HIV-1 mortality rate was estimated to be 1.83 times higher than the HIV-2 mortality rate (acceleration factor $\theta=1.83$, [CI 1.45-2.31]. For direct visual comparison of the shape of Kaplan-Meier and Weibull model curves, survival times of the HIV-1 infected individuals were multiplied by θ (Figure 2D). The similar shapes indicate a highly similar disease course in HIV-1 and HIV-2 infected individuals albeit with lower mortality rates in HIV-2 infection. The corresponding analyses of time to AIDS showed similar results (Appendix, p. 19-20).

Two or more CD4% and CD8% levels were available for 85 and 60 HIV-1 and 48 and 39 HIV-2 infected individuals. The mean rate of decline in CD4% was 0.9% (Standard error [SE] $\pm 0.2\%$) for HIV-1 and 0.4% (SE $\pm 0.2\%$) per year for HIV-2 infected individuals ($p=0.028$, likelihood-ratio test [LRT]). The extrapolated mean CD4% level at estimated date of infection was 22.3% (SE $\pm 1.7\%$) for HIV-1 and 28.0% (SE $\pm 1.3\%$) for HIV-2 infected individuals ($p=0.00094$, LRT). In comparison, the mean CD4% level among HIV-negative individuals was 38.0% (SE $\pm 0.5\%$). Analysis of CD8% over time showed a significantly faster increase among HIV-1 (mean 2.7% per year, SE $\pm 0.5\%$) compared with HIV-2 infected individuals (mean 0.7%, SE $\pm 0.3\%$, $p=0.00015$, LRT), indicating a greater systemic immune activation among HIV-1 infected individuals. No significant differences were found in CD8% levels at infection between HIV-1 and HIV-2 infected individuals. Analysis of CD4% levels by sex showed similar rates of decline in men and women among both HIV-1 and HIV-2 infected individuals, respectively. The mean level of CD4% at the estimated date of infection was 21.4% (SE $\pm 1.2\%$) among HIV-1 infected men compared with 25.8% (SE $\pm 2.3\%$) among HIV-1 infected women ($p=0.057$, LRT); and 26.7% (SE $\pm 1.4\%$) among HIV-2 infected men compared with 35.0% (SE $\pm 3.6\%$) for HIV-2 infected women ($p=0.022$, LRT). Overall, the corresponding analyses for absolute T-cell counts showed similar results with a

mean rate of decline in absolute CD4+ T-cell counts of 22.5 per year for HIV-1 and 12.8 per year for HIV-2 infected individuals ($p=0.060$, LRT). The mean absolute CD4+ T-cell levels at estimated date of infection were 447.9 for HIV-1 and 570.5 for HIV-2 infected individuals ($p=0.012$, LRT) (more details on these analyses are given in Appendix, p. 20-22).

To assess if HIV-2 infected individuals may develop clinical AIDS at higher CD4+ T-cell levels than HIV-1 infected individuals⁵, AIDS was defined purely by clinical manifestations according to WHO staging, without using the immunological criteria.¹⁷ The CD4+ T-cell level at clinical AIDS diagnosis was extrapolated by linear regression for each individual because CD4+ T-cell levels were not routinely determined at the time of clinical AIDS diagnosis and because extrapolation may reduce variation caused by specimen age, sampling time, and handling. Since CD4+ T-cell levels at AIDS diagnosis were independent of information on estimated infection date, we also included data from individuals infected with HIV at enrollment in this analysis (Appendix, p. 22-24). The median CD4% at clinical AIDS was 8.2% (IQR 3.0-13.8) for HIV-1 and 18.2% for HIV-2 infected individuals (IQR 7.2-25.4, $p<0.0001$, two-tailed Mann-Whitney U test, Figure 3). No significant differences in CD4% at clinical AIDS were found between men and women in HIV-1 or HIV-2 infection. Analysis of absolute CD4+ T-cell counts showed similar results with approximately twice as high levels at clinical AIDS for HIV-2 infected individuals.

DISCUSSION

In this study, we show that the majority of HIV-2 infected individuals will progress to HIV-related disease and death without treatment. This observation is based on longitudinal data from HIV-1 or HIV-2 infected individuals with an estimated HIV infection date. The incubation time from infection to development of AIDS can last for several years. Lack of infection date can severely bias studies of natural HIV disease progression. Such data is particularly important when comparing HIV-1 and HIV-2 disease progression due to an epidemic shift (around the year of 2000), in which HIV-1 became more common than HIV-2 in HIV-1/2 endemic countries.^{2,3} Most longitudinal HIV-1 and HIV-2 studies have been based on cohorts established in the late 1980's or during the 1990's, and consequently included/followed patients during some time period between the cohort start and today. Given the crossing nature of the HIV-1 and HIV-2 epidemics in this period, lack of estimated infection date will inevitably result in (1) a selection for HIV-1 infected individuals recruited closer to their infection date compared with HIV-2 infected individuals, and (2) selection for HIV-infected individuals with a slower disease progression than average (with a larger effect among HIV-2 than HIV-1 infected populations since the likelihood of inclusion increases with disease-free time). The effect of these two selective forces will be the opposite. However, their respective impact on the true disease progression rate will depend on average enrollment date, distribution of enrollment dates, and length of follow-up. Other parameters that may influence estimates of disease duration are enrollment age, CD4+ T-cell levels or late-stage disease at enrollment. A previous study from rural Guinea-Bissau with up to 19 years of follow-up indicated similar survival estimates among HIV-1 and HIV-2 infected individuals as presented in the current study¹². However, this study was based mainly on female participants and follow-up data without information on infection date, making interpretations and direct comparisons difficult. The only previous report with estimated date

of HIV-2 infection (n=33, with an average person-years of observation of three years) studied time from infection to non-AIDS defining symptoms (two recorded cases) and clinical AIDS (no cases).⁸ Our study is therefore the first to provide estimates on median time to AIDS or AIDS-related mortality for HIV-2 infected individuals from the date of HIV-2 infection.

Although it is well known that HIV-2 can lead to AIDS and HIV-related death, it is often suggested that the majority of HIV-2 infected individuals never develop AIDS and have the same mortality pattern as the general population.^{3,9,23} Indeed, if the HIV-2 infected population constituted a mixture of two populations with distinctly different survival, it would manifest as a bimodal survival curve where mortality at the initial left part of the curve is dominated by deaths in the subpopulation that develops AIDS and die from the disease. The right part would follow essentially the same pattern as the general population with a much lower mortality rate. We found no indication of such pattern in our data. On the contrary, our survival analysis indicated that HIV-2 infected individuals follow a similar survival curve as HIV-1 infected individuals, albeit at a slower rate. However, the number at risk drop off dramatically after 10 years and estimates for the tail of the distribution are uncertain (as reflected in Weibull CI for survival).

One limitation of our study's generalisability is the imbalance in sex distribution. Whereas our cohort is dominated by men, most previous HIV-2 cohort studies have been dominated by women.^{6,8-10,12,20} It was recently suggested that HIV-infected men had a higher mortality rate than HIV-infected women regardless of infecting HIV type.²⁴ Our data support this, and it is possible that such sex difference may have contributed to the perception that most HIV-2 infected individuals become non- or very slow progressors. However, our stratified mortality

analyses indicated a similar difference in survival time between both HIV-1 and HIV-2 infected men, and HIV-1 and HIV-2 infected women.

HIV-2 infected individuals had higher CD4⁺ T-cell levels both early in infection and at clinical AIDS. Given that HIV-1 infection generally results in higher plasma viral loads than HIV-2 infection, it seems plausible that HIV-2 infected individuals more often present with higher CD4⁺ T-cell counts after the acute infection.^{20,25} However, only one adult case of acute HIV-2 infection, retrospectively diagnosed, has been reported in literature, and no comparisons of viral load dynamics between HIV-1 and HIV-2 infection are available.²⁶ Moreover, it is possible that longer periods of partial immunosuppression, caused by both higher initial CD4⁺ T-cell counts following acute HIV-2 infection, and a slower CD4⁺ T-cell decline, increase the risk of acquiring opportunistic infections at higher CD4⁺ T-cell levels in HIV-2 compared with HIV-1 infection. This would be particularly plausible in settings with high pressure of opportunistic infections. Interestingly, Angin *et al.* recently demonstrated a robust capacity to preserve hematopoietic progenitors and thymic activity among HIV-2 infected individuals.²⁷ An effective replenishment of immune cells combined with a consistent immune activation could explain the higher CD4⁺ T-cell levels observed at AIDS in HIV-2 infection.^{28,29} Differences between HIV-1 and HIV-2 infection are likely related to a lower replication rate of HIV-2, attributed to virus properties, lower immune hyperactivation, better immune control, and/or stronger suppressive effects from the immune system.^{2,30} In the absence of commercial HIV-2 RNA assays, HIV-2 infected patients are generally monitored by CD4⁺ T-cell levels, particularly in endemic areas. Consequently, viral load measurements have not been included as a standard procedure in Guinea-Bissau, preventing us from a comprehensive analysis of viral load data in this study.¹⁹ However, plasma viral loads have been shown to be approximately 10-100 times lower in HIV-2 compared with HIV-1 infected

individuals when matched for CD4 count.²⁵ Interestingly, we found that CD4+ T-cell levels at infection were less variable than with CD4+ T-cell decline between individuals, indicating a stronger association between initial CD4+ T-cell levels and disease progression rate. This is in line with previous reports suggesting that the CD4+ T-cell level following early HIV-1 infection may be a better determinant of disease progression than rates of CD4+ T-cell decline.^{19,31} Disentangling how immune responses differ between HIV-1 and HIV-2 infection may have implications in developing future HIV interventions.

Our results support the recent WHO antiretroviral treatment guidelines of early treatment initiation for all HIV-infected individuals, not only those infected by HIV-1 (<http://www.who.int/hiv/pub/guidelines/earlyrelease-arv/en/>). Moreover, clinical trial data is urgently needed to establish the evidence base for optimal usage of ART in HIV-2 infection.^{1,32}

REFERENCES

1. Gottlieb GS, Eholie SP, Nkengasong JN, et al. A call for randomized controlled trials of antiretroviral therapy for HIV-2 infection in West Africa. *Aids* 2008; **22**(16): 2069-72; discussion 73-4.
2. de Silva TI, Cotten M, Rowland-Jones SL. HIV-2: the forgotten AIDS virus. *Trends Microbiol* 2008; **16**(12): 588-95.
3. Tchounga B, Ekouevi DK, Balestre E, Dabis F. Mortality and survival patterns of people living with HIV-2. *Current opinion in HIV and AIDS* 2016; **11**(5): 537-44.
4. De Cock KM, Adjuorlolo G, Ekpini E, et al. Epidemiology and transmission of HIV-2. Why there is no HIV-2 pandemic. *Jama* 1993; **270**(17): 2083-6.
5. Martinez-Steele E, Awasana AA, Corrah T, et al. Is HIV-2- induced AIDS different from HIV-1-associated AIDS? Data from a West African clinic. *Aids* 2007; **21**(3): 317-24.
6. Drylewicz J, Matheron S, Lazaro E, et al. Comparison of viro-immunological marker changes between HIV-1 and HIV-2-infected patients in France. *Aids* 2008; **22**(4): 457-68.
7. Kanki PJ, Travers KU, S MB, et al. Slower heterosexual spread of HIV-2 than HIV-1. *Lancet* 1994; **343**(8903): 943-6.
8. Marlink R, Kanki P, Thior I, et al. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science* 1994; **265**(5178): 1587-90.
9. Poulsen AG, Aaby P, Larsen O, et al. 9-year HIV-2-associated mortality in an urban community in Bissau, west Africa. *Lancet* 1997; **349**(9056): 911-4.
10. Hansmann A, Schim van der Loeff MF, Kaye S, et al. Baseline plasma viral load and CD4 cell percentage predict survival in HIV-1- and HIV-2-infected women in a community-based cohort in The Gambia. *Journal of acquired immune deficiency syndromes* 2005; **38**(3): 335-41.
11. Holmgren B, da Silva Z, Vastrup P, et al. Mortality associated with HIV-1, HIV-2, and HTLV-I single and dual infections in a middle-aged and older population in Guinea-Bissau. *Retrovirology* 2007; **4**: 85.
12. van Tienen C, Schim van der Loeff M, Peterson I, et al. HTLV-1 and HIV-2 infection are associated with increased mortality in a rural West African community. *PloS one* 2011; **6**(12): e29026.
13. Alabi AS, Jaffar S, Ariyoshi K, et al. Plasma viral load, CD4 cell percentage, HLA and survival of HIV-1, HIV-2, and dually infected Gambian patients. *Aids* 2003; **17**(10): 1513-20.
14. Europe CECA-CAoStAaDi. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Concerted Action on SeroConversion to AIDS and Death in Europe. *Lancet* 2000; **355**(9210): 1131-7.
15. Gilleece Y, Chadwick DR, Breuer J, et al. British HIV Association guidelines for antiretroviral treatment of HIV-2-positive individuals 2010. *HIV medicine* 2010; **11**(10): 611-9.
16. Rowland-Jones SL, Whittle HC. Out of Africa: what can we learn from HIV-2 about protective immunity to HIV-1? *Nature immunology* 2007; **8**(4): 329-31.
17. WHO. WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children. 2007. <http://www.who.int/hiv/pub/guidelines/HIVstaging150307.pdf>.
18. Esbjornsson J, Mansson F, Kvist A, et al. Increased survival among HIV-1 and HIV-2 dual-infected individuals compared to HIV-1 single-infected individuals. *Aids* 2014; **28**(7): 949-57.

19. Esbjornsson J, Mansson F, Kvist A, et al. Inhibition of HIV-1 disease progression by contemporaneous HIV-2 infection. *The New England journal of medicine* 2012; **367**(3): 224-32.
20. van der Loeff MF, Larke N, Kaye S, et al. Undetectable plasma viral load predicts normal survival in HIV-2-infected people in a West African village. *Retrovirology* 2010; **7**: 46.
21. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982; **38**(4): 933-42.
22. Deeks SG, Walker BD. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity* 2007; **27**(3): 406-16.
23. de Silva TI, Peng Y, Leligidowicz A, et al. Correlates of T-cell-mediated viral control and phenotype of CD8(+) T cells in HIV-2, a naturally contained human retroviral infection. *Blood* 2013; **121**(21): 4330-9.
24. Jespersen S, Hønge BL, Esbjornsson J, et al. Differential effects of sex in a West African cohort of HIV-1, HIV-2 and HIV-1/2 dually infected patients: men are worse off. *Tropical medicine & international health : TM & IH* 2016; **21**(2): 253-62.
25. Andersson S, Norrgren H, da Silva Z, et al. Plasma viral load in HIV-1 and HIV-2 singly and dually infected individuals in Guinea-Bissau, West Africa: significantly lower plasma virus set point in HIV-2 infection than in HIV-1 infection. *Archives of internal medicine* 2000; **160**(21): 3286-93.
26. Besnier JM, Barin F, Baillou A, Liard F, Choutet P, Goudeau A. Symptomatic HIV-2 primary infection. *Lancet* 1990; **335**(8692): 798.
27. Angin M, Wong G, Papagno L, et al. Preservation of Lymphopoietic Potential and Virus Suppressive Capacity by CD8+ T Cells in HIV-2-Infected Controllers. *Journal of immunology* 2016; **197**(7): 2787-95.
28. Nyamweya S, Townend J, Zaman A, et al. Are plasma biomarkers of immune activation predictive of HIV progression: a longitudinal comparison and analyses in HIV-1 and HIV-2 infections? *PloS one* 2012; **7**(9): e44411.
29. Buggert M, Frederiksen J, Lund O, et al. CD4+ T cells with an activated and exhausted phenotype distinguish immunodeficiency during aviremic HIV-2 infection. *Aids* 2016; **30**(16): 2415-26.
30. MacNeil A, Sarr AD, Sankale JL, Meloni ST, Mboup S, Kanki P. Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (HIV-2) infection than in HIV-1 infection. *J Virol* 2007; **81**(10): 5325-30.
31. Audige A, Taffe P, Rickenbach M, et al. Low postseroconversion CD4 count and rapid decrease of CD4 density identify HIV+ fast progressors. *AIDS research and human retroviruses* 2010; **26**(9): 997-1005.
32. Matheron S, Descamps D, Gallien S, et al. First line raltegravir/emtricitabine/tenofovir combination in HIV-2 infection: phase 2 non-comparative trial (ANRS 159 HIV-2). *Clin Infect Dis* 2018.

Acknowledgements

We would like to thank all participants in the study. We also thank all the personnel at the Health Station of the 2a esquadra, Bissau, and at the National Public Health Laboratory in Bissau. The study was supported by the Department for Research Cooperation (SAREC) at the Swedish International Development Agency (SIDA), the Swedish Research Council (No. 350-2012-6628 and 2016-01417 for JE; No. 2016-02285 for MJ; No. 321-2012-3274 for PM), Swedish Society of Medical Research, Medical Faculty at Lund University (ALF), and Region Skåne FoU.

Author contributions

J.E. and F.M. contributed equally to this study. J.E., F.M., P.M. and H.N. interpreted the data and were responsible for the overall study design. E.M.F., M.J., P.M., and H.N. were responsible for the overall project coordination. F.M., H.N., A.J.B. were medically and organisationally responsible for the clinical sites with biological samples of the study participants in the cohort. S.A. established the diagnostic strategies, and Z.J.d.S. was responsible for analyses of HIV serology and T-cell count determinations at the laboratory in Guinea-Bissau. F.M. and H.N. coordinated the laboratory and clinical work, performed database entry and cleaning. J.L.L. coordinated clinical follow-up and mortality reports. A.P. coordinated cytometric follow-up. J.E. and A.K. analysed the data. J.E., A.K. and P.E.I. contributed in statistical analyses. M.J. and S.R-J. participated in interpretation of the results. J.E. wrote the manuscript. All authors read and approved the manuscript.

Conflicts of interest

The authors or their institutions declare no competing financial interests, and did not at any time receive payment or services from a third party (government, commercial, private

foundation, etc.) for any aspect of the submitted work (including data monitoring board, study design, manuscript preparation, statistical analysis, etc.). The authors have no patents, whether planned, pending or issued, broadly relevant to the work.

Correspondence

Correspondence and requests for materials should be addressed to J.E. (joakim.esbjornsson@med.lu.se).

Members of the SWEGUB CORE group

Current members of the Sweden Guinea-Bissau Cohort Research (SWEGUB CORE) group includes Babetida N'Buna, Antonio J. Biague, Ansu Biai, Cidia Camara, Joakim Esbjörnsson, Marianne Jansson, Sara Karlson, Jacob Lindman, Patrik Medstrand, Fredrik Månsson, Hans Norrgren, Angelica A. Palm, Gülsen Özkaya Sahin, Zacarias José da Silva, and Sten Wilhelmson.

TABLES

Table 1. Characteristics of the study groups with estimated date of HIV-infection or that remained HIV-negative during follow-up¹.

	HIV-1	HIV-2	HIV-negative
Number of individuals	225	87	2984
Number of men (%) ²	188 (84%)	72 (83%)	2679 (90%)
Median age at estimated HIV-infection (years, IQR) ³	36 (30-45)	38 (31-45)	35 (27-44) ⁹
Median follow-up time (years, IQR) ⁴	8.2 (3.8-13.5)	10.8 (4.9-16.2)	6.4 (3.9-14.3) ⁹
Median time between last HIV-negative and first HIV-positive sample (years, IQR)	2.7 (1.5-4.1)	2.1 (1.2-3.9)	Not applicable
Recorded deaths (%) ⁵	104 (46%)	26 (30%)	276 (9%)
Number of censored cases (%) ⁵	121 (54%)	61 (70%)	2708 (91%)
Median time to death or censoring (years, IQR) ⁶	5.5 (3.1-8.3)	8.0 (4.4-13.5)	6.0 (3.7-13.6) ⁹
Median age at death (years, IQR) ⁷	51 (42-60)	63 (52-64)	74 (63-83)
Number of individuals developing AIDS (%)	121 (54%)	37 (43%)	Not applicable
Median time to AIDS or censoring (years, IQR) ⁸	4.2 (2.3-6.9)	6.5 (3.5-11.8)	Not applicable

¹There were no trends or significant pairwise differences between groups except as noted (significance cut-off: $p < 0.05$). Correction for multiple comparisons according to Holm-Šidák. IQR: Interquartile range.

²HIV-1 vs. HIV-negatives: $p = 0.0067$, HIV-2 vs. HIV-negatives: $p = 0.048$, two-tailed Fisher's exact test (FET).

³HIV-1 vs. HIV-negatives: $p < 0.0001$; HIV-2 vs. HIV-negatives: $p = 0.068$, two-tailed Mann-Whitney U test (M-W).

⁴HIV-2 vs. HIV-negatives: $p = 0.00092$, Log-Rank test as estimated by reverse Kaplan-Meier estimator.

⁵HIV-1 vs. HIV-2: $p = 0.0103$; HIV-1 vs. HIV-negatives: $p < 0.0001$; HIV-2 vs. HIV-negatives: $p < 0.0001$, FET.

⁶HIV-1 vs. HIV-2: $p < 0.0001$; HIV-1 vs. HIV-negatives: $p < 0.0001$, M-W.

⁷HIV-1 vs. HIV-2: $p = 0.00022$; HIV-1 vs. HIV-negatives: $p < 0.0001$; HIV-2 vs. HIV-negatives: $p < 0.0001$, Log-Rank test as estimated by Kaplan-Meier analysis.

⁸HIV-1 vs. HIV-2: $p < 0.0001$, M-W.

⁹As estimated from or at enrollment for HIV-negatives.

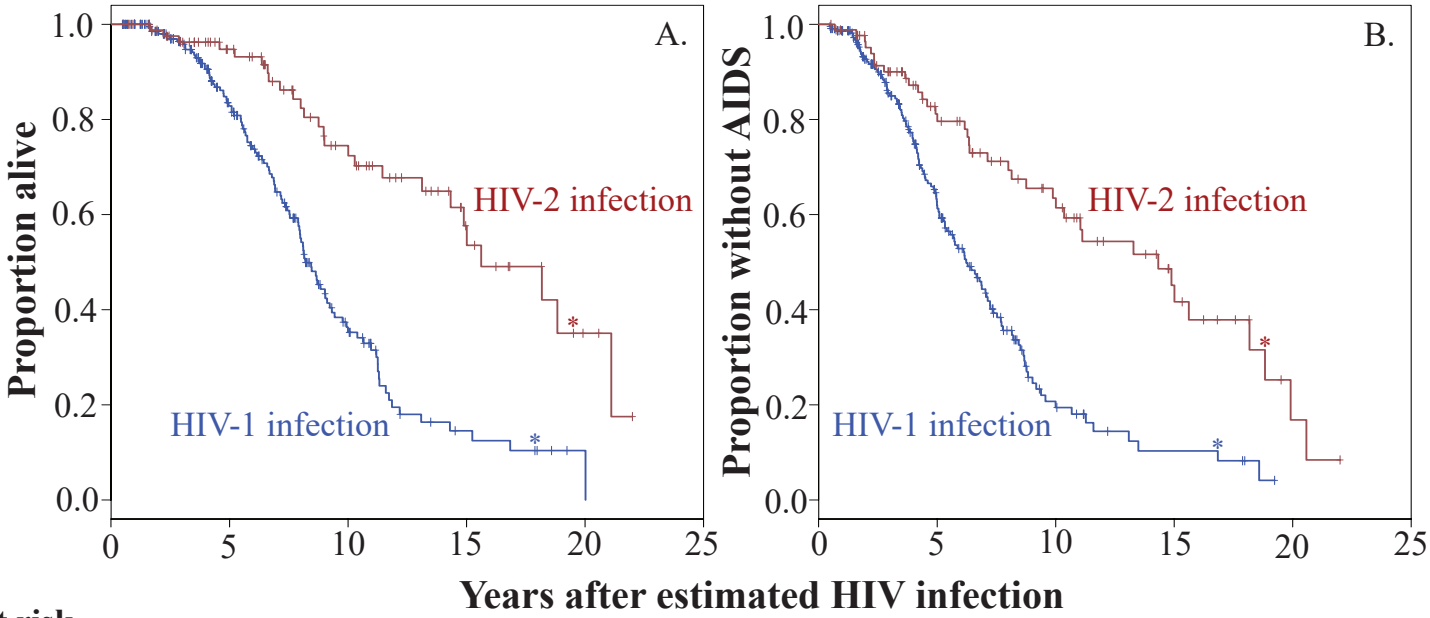
FIGURE LEGENDS

Figure 1. Kaplan-Meier curves of survival (A) and AIDS-free time (B) of HIV-1 and HIV-2 infected individuals. Tick marks indicate participants with censored data. Asterisks indicate the time-point in each group when five participants are still at mortality risk and risk of developing AIDS. The numbers of individuals at risk are given below the figures at 5-year intervals.

Figure 2. Survival characteristics in HIV-1 and HIV-2 infection. (A) Comparisons of survival curves of HIV-1 (blue) and HIV-2 (red) infected individuals and the expected survival of matched HIV-negative individuals. (B) Weibull estimates of the survival function (solid lines) with 95% confidence intervals (dashed lines) for HIV-1 (blue) and HIV-2 (red) infected individuals. Kaplan-Meier curves are plotted for comparison. (C) Probability density functions of the Weibull distributions showing the relative likelihood of mortality in the infected population over time. (D) Kaplan-Meier and Weibull curves (with 95% confidence intervals, dashed lines) for HIV-2 infected individuals (red) compared with a simulated HIV-2 dataset, generated by multiplying the survival times of the HIV-1 infected individuals by the acceleration factor θ determined by comparing HIV-1 and HIV-2 infected individuals in a Weibull regression model (blue).

Figure 3. Pirate plots of estimated CD4% levels at clinical AIDS for HIV-1 (blue) and HIV-2 (red) infected individuals. Clinical AIDS was determined according to the World Health Organization clinical staging system¹¹. Pirate plots combining individual observations (black points), box plots of the median, interquartile range and 95% quantile range, and a smoothed density distributions of the data are shown for each group. Comparison between groups was performed by the two-tailed Mann-Whitney U Test.

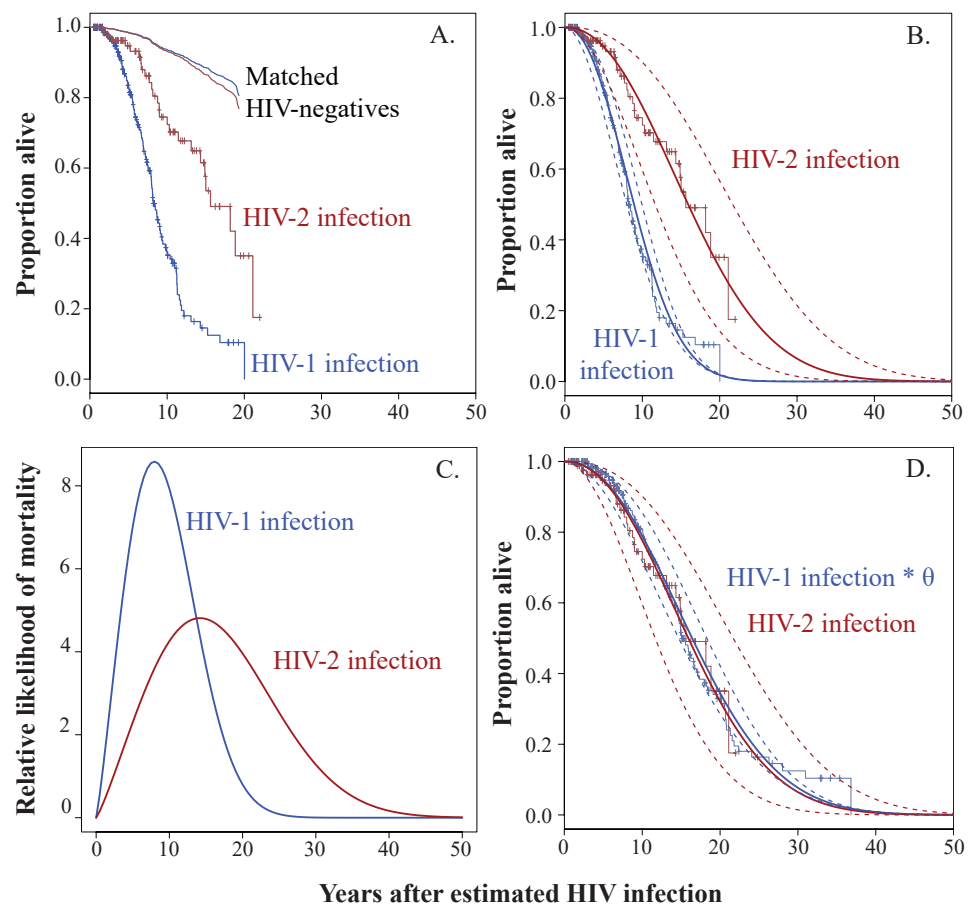
Figure_1



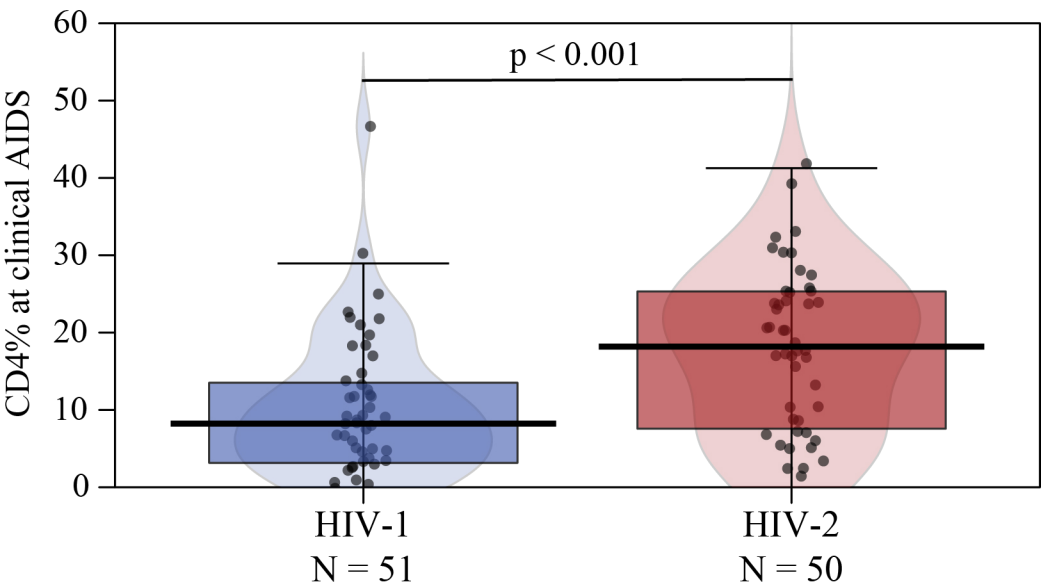
Number at risk

HIV-1 infection	225	125	33	7	1	0	225	93	15	5	0	0
HIV-2 infection	87	61	35	14	3	0	87	52	31	13	2	0

Figure_2



Figure_3



Appendix

Web extra material

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Esbjörnsson J, Månsson F, Kvist A, et al. Long follow-up reveals high proportions of HIV-2 related AIDS and mortality: an open prospective cohort study

Supplementary Methods

1. Serological analysis and T-cell analysis
2. Statistics

Supplementary Results with Supplementary Tables and Figures

1. General description of the cohort and analyses of follow-up
2. Analyses of rate of follow-up
3. Analyses of loss to follow-up
4. Detailed description of HIV seroincident individuals lost to follow-up
5. Uncertainty in estimated HIV seroconversion date
6. HIV-related mortality and time to WHO 4 stage
7. Potential bias of antiretroviral treatment initiation and HIV-1 and HIV-2 dual-infection
8. Survival analysis stratified by sex
9. Characteristics of participants HIV-infected at enrollment
10. Trajectory of time to AIDS in HIV-1 and HIV-2 infection
11. Analysis of T-cell dynamics
12. CD4⁺ T-cell levels at AIDS

Supplementary References

Supplementary Methods

1. Serological analysis and T-cell analysis

In 1990-1994, sera were screened for HIV-1 and HIV-2 antibodies by enzyme-linked immunosorbent assay (ELISA) with the use of the Behring anti-HIV-1/HIV-2 (Behring, Marburg, Germany) and/or Wellcozyme recombinant anti-HIV-1 (Wellcome, Dartford, UK) and an in-house HIV-2 (SBL6669) ELISA assay. From 1995 and onwards, the ELISA test for screening was Behring Enzygnost HIV-1/HIV-2 Plus ELISA (Behring). Confirmation of positive results was done with western blot analysis (Diagnostic Biotechnology anti-HIV-1 blot 2.2, Science park, Singapore, or in-house anti-HIV-2) and dually HIV-1/HIV-2-positive samples were confirmed by Pepti-lav (Sanofi Diagnostics Pasteur, Marnes-la-Coquette, France) in the years 1990-1997¹. In 1996, an alternative confirmation strategy was introduced using Capillus HIV-1/HIV-2 (Cambridge Biotech Limited, Galway, Ireland) in parallel with the Behring Enzygnost HIV-1/HIV-2 Plus ELISA and Immunocomb II HIV-1 and 2 BiSpot RST (Orgenics, Yavne, Israel) for differentiation between HIV-1 and HIV-2². The used screening assays have all been evaluated in parallel in order to avoid different performance characteristics.

T-lymphocyte subsets were determined at the National Public Health Laboratory (LNSP), by conventional flow cytometry (Until 2005: FACStrak; Becton Dickinson, San Jose, CA, and from 2006 and onwards: CyFlow, Partec, Münster, Germany). Leukocyte counts were performed with a cell counter until 2005 (Coulter Counter CBC5; Coulter Electronics Ltd, Luton, England), and by CyFlow from 2006 and onwards. Since viral load (VL) measurements have not been included as a standard procedure in Guinea-Bissau, we were unable to include VL data in our analyses. Due to limitations in sample size and sample availability, we were not able to do any retrospective and valid sequential VL measurements of the samples included in this study.

2. Statistics

Except for the statistical methods presented in the main text, the two-tailed Fisher's exact test was used to analyse categorical data and the two-tailed Mann-Whitney U test was used to analyse continuous data. Correction for multiple comparison was performed by the Holm-Šidák procedure³. The above analyses, Kaplan-Meier analyses, Cox regressions, linear regressions, and linear mixed models were performed in IBM® SPSS® Statistics version 23 (IBM Corporation), whereas the expected survival analysis, Weibull analysis, and Pirate plots were done in R 3.2.2 using the packages survival, flexsurv, relsurv, and yarr⁴. Pirate plots combine individual observations (points), box plots of the median, interquartile range and 95% quantile range, and a smoothed density distribution of the data for each group. These plots were introduced for the first time in 2015 and are more informative compared with common bar and box plots (that can mask important patterns in data like multiple modes and skewness).

Supplementary Results with Supplementary Tables and Figures

Analysing data from HIV-1 and HIV-2 infected individuals is accompanied by some potential caveats. Not least when there are clear epidemiological differences between groups over time, including a shift from HIV-2 to HIV-1 being the most common HIV type in the study setting^{5,6}. Here, we present a detailed evaluation of such caveats and their potential impact on our results.

1. General description of the cohort and analyses of follow-up

A prospective open cohort of police officers in Guinea-Bissau was initiated February 6, 1990. Since then, new participants have continuously been included, except for a temporary closure of inclusion from June 1998 until the end of 2002 as a result of the civil war in 1998–1999 (Figure S1). All persons with a regular employment in the Guinea-Bissau police force were eligible for the study. Participation was voluntary, and >98% accepted to be included. The study was approved by the ethical committees of the Ministry of Health in Guinea-Bissau, and at Lund University and Karolinska Institute, Sweden. All participants received detailed information about the study before inclusion, as well as pre- and post-HIV counselling. Up to 2011, only oral informed consent was required. Since then, written informed consent has been required from all study participants. All participants that tested HIV positive during the study period have been informed about antiretroviral treatment (ART), the national treatment guidelines (which have followed WHO recommendations throughout the study period), and been offered ART accordingly. However, it is important to emphasize that the WHO recommendations have been modified between 1990–2013. Between 2006 and 2010 it was recommended that HIV-infected patients with CD4+ T-cell counts below 200 should be offered ART, and between 2011 and 2013, HIV-infected patients with CD4+ T-cell counts below 350 were offered ART. Moreover, our team has actively followed the participants in the police cohort, and the vast majority of eligible participants have been offered ART during the study period (the few that were not offered ART were mainly participants that were lost-to-follow-up, see below for a detailed discussion about loss-to-follow-up). Overall, 872 participants were tested HIV-positive during the study period, of which 74 participants eligible for ART (8%) remained untreated during follow-up. Of the 74 participants, 32 developed WHO stage 3 symptoms, 15 developed WHO stage 4 symptoms, and 14 died of HIV-related illnesses during the study period. Moreover, of the 74 participants, nineteen participants declined to start treatment even after detailed information about the risks not taking ART, and of these 3 died during follow-up. Importantly, in those cases where ART were offered and declined, continuous support has been given together with motivating information, and eventually many started up treatment. Of the remaining 55 participants: two participants started ART but chose to stop treatment only few weeks after treatment initiation and were therefore considered untreated; 48 participants reached the eligibility criteria at their last visit close to study closure and had therefore not initiated ART at that time; and for five participants information on whether they were offered ART or not was missing.

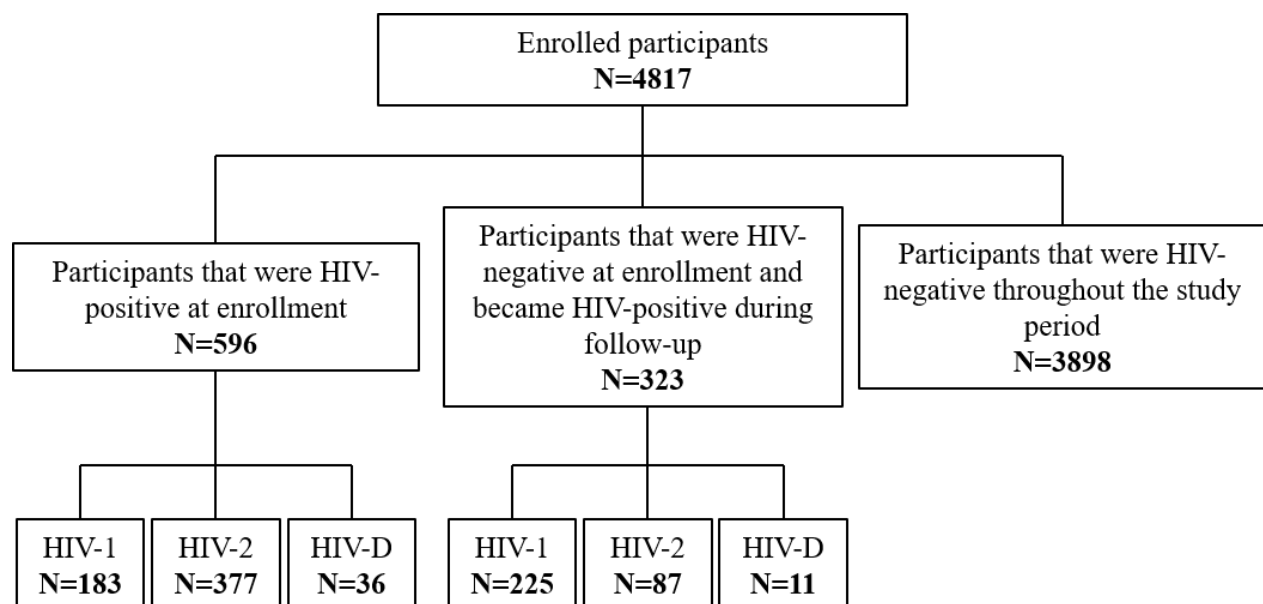


Figure S1. Flow chart over the enrolled participants during the study period (1990-2013). HIV-1: HIV-1 infected participants; HIV-2: HIV-2 infected participants; HIV-D: HIV-1 and HIV-2 dual-infected participants.

Blood samples for serology and CD4⁺ T-cell counts were collected at inclusion and at follow-up visits scheduled at intervals of 12-18 months except for a temporary closure of follow-up visits from June 1998 and July 2000 due to the civil war. This means that retesting and confirmation of previous HIV-testing was done at every follow-up visit for both HIV negative and HIV positive participants. Inclusion continued until September, 2009, when 4817 participants had been included in the cohort. General follow-up visits were performed until February, 2011. However, follow-up of the HIV-1 and HIV-2 positive individuals infected after enrollment continued until September 2013. Causes of death were ascertained by staff at the health care unit of the Ministry of Interior in Guinea-Bissau after cross-checking with reports of deceased employees at the ministry. The cause of death, if unknown by the health care unit, was determined by information from family members and/or co-workers. If cause of death could not be identified it was included in the non-HIV-related category (5% of all [including both individuals with an estimated HIV infection date, and individuals that were HIV-infected at enrollment] HIV-1 and 11% of HIV-2 single infected individuals died of non-HIV-related causes).

AIDS was defined according to WHO staging as CD4⁺ T-cell counts ≤ 200 cells per μL or CD4% counts $\leq 14\%$ (immunological criteria), or AIDS-defining symptoms (WHO 4 clinical criteria, <http://www.who.int/hiv/pub/guidelines/HIVstaging150307.pdf>). Of the 872 participants that tested HIV-positive, 408 participants tested positive for HIV-1 and 464 participants tested positive for HIV-2. Among the HIV-1 positive individuals, 213 (52%) never developed AIDS during follow-up, 70 (17%) developed AIDS based solely on immunological criteria, 58 (14%) developed AIDS based solely on WHO 4 criteria, and 67 (16%) developed AIDS fulfilling both the immunological and WHO 4 criteria. Among the HIV-2 positive individuals, 295 (64%) never developed AIDS during follow-up, 65 (14%) developed AIDS based solely on immunological criteria, 66 (14%) developed AIDS based solely on WHO 4 criteria, and 38 (8%) developed AIDS fulfilling both the immunological and WHO 4 criteria. The proportion of different WHO 4 defining symptoms were similar between HIV-1 and HIV-2 infected individuals, with the most common symptoms being: Unexplained involuntary weight loss (100% of HIV-1 WHO 4 cases, and 100% of HIV-2 WHO 4 cases), fever (90% of HIV-1 WHO 4 cases; and 83% of HIV-2 WHO 4 cases); and diarrhoea (35% of HIV-1 WHO 4 cases, and 48% of HIV-2 WHO 4 cases).

In total, 225 HIV-1 and 87 HIV-2 seroincident individuals were included in the present study. The HIV seroincident individuals have by definition been tested negative for HIV at inclusion and then positive for HIV sometime during follow-up. This means that all seroincident individuals have attended at least one follow-up visit. Four participants, with ART-periods shorter than four months before death, were not censored for mortality at ART initiation as the ART was considered too short to have affected the AIDS-related deaths. Individuals with recorded seroconversion from single HIV-positive status to dual HIV-1 and HIV-2 reactivity were censored from the analyses at the day of last serological single HIV-positive status (N=4).

We used the reverse Kaplan-Meier (KM) method to estimate median follow-up times. This method has been described to be a robust measure of the follow-up length^{7,8}. The rationale behind the reverse KM analysis is that we cannot know how long an individual that developed an event would have been followed if had it not developed the event. Hence, the censoring becomes the event and events become censoring in the reverse KM analysis⁷. In our cohort, 3761 individuals (78%) had at least one follow-up visit and the median follow-up time among these individuals was 7.4 years (Table S1). Since police officers were included into the cohort continuously between 1990 and 2009, we also stratified the cohort by inclusion date (Table S1).

Table S1. Number of individuals enrolled, with at least one follow-up visit and reverse Kaplan-Meier estimates of median follow-up times for the entire cohort. Number of individuals enrolled and with at least one follow-up (FU) visit (percentage of enrolled in parenthesis). Reverse Kaplan Meier estimates of the median follow-up time among individuals with at least one follow-up visit (Median FU time). The cohort was stratified by inclusion date.

Inclusion period	Enrolled	≥ 1 FU visit (%)	Median FU time (years) (95% CI)
1990-1994	1984	1798 (90%)	14.7 (14.4-15.0)

1995-1999	820	670 (82%)	10.1 (9.6-10.6)
2000-2004	1234	936 (76%)	5.4 (5.2-5.6)
2005-2009	779	357 (46%)	3.3 (3.3-3.4)
Total	4817	3761 (78%)	7.4 (6.8-8.0)

The corresponding analyses for the HIV seroincident individuals showed long median follow-up times from inclusion for both groups (Table S2). As expected, the median follow-up times for the seroincident individuals were longer than for the entire cohort (the longer an individual is observed, the greater the cumulative likelihood that it will seroconvert under the observation time). This analysis also showed that the majority of these individuals (67%) were included in the cohort during 1990-1994.

Table S2. Reverse Kaplan-Meier estimates of median follow-up time since inclusion for HIV seroincident individuals. Number of HIV seroincident cases (N) and reverse Kaplan Meier estimates of the median follow-up time since inclusion (Median FU time). Data is stratified by inclusion date. P-values from Log-Rank (LR) test comparing median FU time between HIV-1 and HIV-2 infected individuals are presented.

Inclusion period	N		Median FU time (years, 95% CI)		p (LR)
	HIV-1	HIV-2	HIV-1	HIV-2	
1990-1994	134	74	18.8 (17.0-20.6)	19.8 (17.0-20.6)	0.82
1995-1999	31	8	14.6 (13.5-15.7)	9.8 (0.3-19.3)	0.55
2000-2004	44	3	5.7 (5.3-6.1)	5.8 (4.3-7.4)	0.62
2005-2009	16	2	4.7 (4.1-5.3)	4.8 (NA ¹)	0.78
Total	225	87	16.2 (14.6-17.8)	18.8 (17.0-20.7)	0.034

¹NA: Not applicable

A more relevant measure of follow-up time for the seroincident individuals would be from the date of seroconversion (Table S3). This is the time origin (time zero) for all the analyses we present in the main text, and it is from this point the time to mortality and AIDS is measured. To keep the number of individuals in each section reasonably large, we stratify the seroincident individuals based on the date of seroconversion over only two equally long time periods (1990-1999 and 2000-2009). There were no significant differences in median follow-up time between the HIV-1 and HIV-2 infected individuals, either in total or within each time period.

Table S3. Reverse Kaplan-Meier estimates of median follow-up time since seroconversion for HIV seroincident individuals. Number of HIV seroincident cases (N), reverse Kaplan Meier estimates of the median follow-up time since seroconversion (Median FU time). Data is stratified by HIV seroconversion date. P-values from Log-Rank (LR) test comparing median FU time between HIV-1 and HIV-2 infected individuals are presented.

HIV seroconversion period	N		Median FU time since seroconversion (years, 95% CI)		p (LR)
	HIV-1	HIV-2	HIV-1	HIV-2	
1990-1999	90	61	14.5 (8.5-20.5)	13.8 (9.5-18.0)	0.36
2000-2009	135	26	5.2 (3.8-6.6)	6.5 (3.8-9.2)	0.11
Total	225	87	8.2 (7.0-9.5)	10.8 (8.3-13.2)	0.06

Another way of measuring follow-up time can be time under observation (TUO). However, it is important to remember that time under observation systematically underestimates the potential follow-up time of subjects, and increasingly so, the greater the event hazard. Thus, for two groups with identical follow-up but with different hazard, time under observation will take on lower values for the group with the greater hazard⁷. The follow-up times, estimated as time from inclusion to last follow-up visit, for all individuals in the cohort with at least one follow-up visit are presented in Table S4.

Table S4. Intervals of time under observation from inclusion in the entire cohort. The number of individuals in the cohort with at least one follow-up visit stratified by time under observation and inclusion date. Time under observation was estimated as the time from inclusion to the last follow-up visit.

Inclusion period	>3 years TUO (%)	>5 years TUO (%)	>10 years TUO (%)	>15 years TUO (%)
1990-1994	1591 (88%)	1429 (79%)	1133 (63%)	660 (37%)
1995-1999	530 (79%)	500 (75%)	302 (45%)	21 (3%)
2000-2004	694 (74%)	489 (52%)	0 (0%)	NA ¹
2005-2009	217 (61%)	15 (4%)	NA ¹	NA ¹
Total	3032 (81%)	2433 (65%)	1435 (38%)	681 (18%)

¹NA: Not applicable

We also analysed the times under observation for the HIV-1 and HIV-2 infected individuals from inclusion (Table S5). The proportion of individuals observed over different intervals of time was similar between groups both in total and within each inclusion period, except for an overall difference in proportions of HIV-1 and HIV-2 infected individuals with a longer time under observation (TUO) than 15 years. This is likely explained by the faster disease progression observed in HIV-1 compared with HIV-2 infected individuals.

Table S5. Intervals of time under observation from inclusion to last follow-up visit for HIV seroincident individuals. The number of HIV seroincident individuals stratified by inclusion date and time under observation (TUO). Time under observation was estimated as the time from inclusion to last follow-up visit. Two-tailed Fisher's exact test was used to test if there was any significant difference between the HIV-1 and HIV-2 groups.

Inclusion period	Type	>3y TUO	>5y TUO	>10y TUO	>15y TUO
1990-1994	HIV-1	133/134 (99%)	130/134 (97%)	111/134 (83%)	61/134 (46%)
	HIV-2	72/74 (97%)	69/74 (93%)	58/74 (78%)	41/74 (55%)
	p-value	p=0.29	p=0.29	p=0.46	p=0.19
1995-1999	HIV-1	30/31 (97%)	30/31 (97%)	24/31 (77%)	7/31 (23%)
	HIV-2	6/8 (75%)	6/8 (75%)	6/8 (75%)	2/8 (25%)
	p-value	p=0.10	p=0.10	p=1.0	p=1.0
2000-2004	HIV-1	36/44 (82%)	28/44 (64%)	0/44 (0%)	NA ¹
	HIV-2	3/3 (100%)	2/3 (67%)	0/3 (0%)	NA ¹
	p-value	p=1.0	p=1.0	p=1.0	NA ¹
2005-2009	HIV-1	14/16 (88%)	5/16 (31%)	NA ¹	NA ¹
	HIV-2	2/2 (100%)	1/2 (50%)	NA ¹	NA ¹
	p-value	p=1.0	p=1.0	NA ¹	NA ¹
Total	HIV-1	213/225 (95%)	193/225 (86%)	135/225 (60%)	68/225 (30%)
	HIV-2	83/87 (95%)	78/87 (90%)	62/87 (71%)	43/87 (49%)
	p-value	p=1.0	p=0.46	p=0.068	p=0.0023

¹NA: Not applicable

Next, we investigated the times under observation for the HIV seroincident individuals from the time of seroconversion (Table S6). This analysis showed a generally larger difference in proportion of HIV-2 infected individuals with long TUO compared with HIV-1 infected individuals. As would be expected when comparing two groups with different mortality rates, this difference became gradually larger with longer TUO reflecting the generally longer progression-time seen among HIV-2 infected individuals. The reason why this was not as clear in the analyses of time under observation from inclusion is likely explained by differences in the HIV-1 and HIV-2 epidemics (although there was a tendency to a significant difference in the higher TUO groups in the total analysis,

Table S5). It has been shown in several studies (both by us and others) that the prevalence of HIV-1 has been increasing in West Africa since the beginning of the 1990's (even though the prevalence of HIV-1 seems to be levelling off the last five years). The opposite pattern has been shown for HIV-2. These epidemiological patterns are true also for Guinea-Bissau. This explains why most of the HIV-2 infected individuals were included early in the cohort, while the inclusions of HIV-1 infected individuals were more evenly distributed between the two stratified periods (Table S6). The median year of HIV infection were 2000 for the HIV-1 infected individuals and 1996 for the HIV-2 infected individuals. This likely explains why there in general were no significant differences between the groups in time under observation from inclusion in the cohort.

Table S6. Intervals of time under observation from seroconversion for HIV seroincident individuals. The number of HIV seroincident individuals stratified by seroconversion date and time under observation (TUO). Time under observation was estimated as the estimated time of infection to last follow-up visit. Two-tailed Fisher's exact test was used to test if there was any significant difference between the HIV-1 and HIV-2 groups in each interval.

HIV seroconversion period		>3y TUO	>5y TUO	>10y TUO	>15y TUO
1990-1999	HIV-1	85/90 (94%)	71/90 (79%)	24/90 (27%)	7/90 (8%)
	HIV-2	51/61 (84%)	45/61 (74%)	29/61 (48%)	14/61 (23%)
	p-value	p=0.049	p=0.56	p=0.0097	p=0.015
2000-2009	HIV-1	87/135 (64%)	54/135 (40%)	9/135 (7%)	NA ¹
	HIV-2	23/26 (88%)	16/26 (62%)	6/26 (23%)	NA ¹
	p-value	p=0.020	p=0.053	p=0.018	NA ¹
Total	HIV-1	172/225 (76%)	125/225 (56%)	33/225 (15%)	7/225 (3%)
	HIV-2	74/87 (85%)	61/87 (70%)	35/87 (40%)	14/87 (16%)
	p-value	p=0.12	p=0.021	p<0.0001	p<0.0001

¹NA: Not applicable

As described above, the prevalence of HIV-2 in Guinea-Bissau and West Africa has declined since the 1990s. Therefore, the proportion of HIV-2 infected individuals is larger among individuals included early in the cohort than at later inclusion periods (Table S2). This could lead to a skew in follow-up time between the HIV-1 and HIV-2 groups, since individuals included early have a greater chance of being followed for a longer time. To investigate if this could have affected our conclusions, we conducted a Kaplan-Meier analysis of study participants that were included into the cohort 1990-1994. As shown above, median follow-up time (Table S2) and time under observation during different time intervals from inclusion (Table S5) are similar for HIV-1 and HIV-2 infected individuals included 1990-1994. 134 HIV-1 and 74 HIV-2 infected individuals were included in the analysis. The median survival time was 8.1 years (95% confidence interval [CI] 7.4-8.8) in HIV-1 and 15.6 years (CI 11.9-19.3) in HIV-2 infected individuals ($p<0.0001$, Log-Rank test). The median time to AIDS was 6.1 years (CI 4.9-7.4) in HIV-1 and 14.9 years (CI 10.5-19.3) in HIV-2 infected individuals ($p<0.0001$, Log-Rank test). These estimates are almost identical to the estimates from the original analysis and the confidence intervals were similar, demonstrating that individuals included late in the cohort and with generally shorter follow-up times have little impact on the results from the analysis. However, the date of inclusion does not carry any information of the infection date. Therefore, we also performed a second Kaplan-Meier analysis, comparing the individuals that seroconverted 1990-1999. This analysis included 90 HIV-1 and 61 HIV-2 infected individuals. The median survival time was 8.7 years (CI 7.6-9.8) in HIV-1 and 18.2 years (CI 12.8-23.6) in HIV-2 infected individuals ($p<0.0001$, Log-Rank test). The median time to AIDS was 6.1 years (95% CI 4.7-7.6) in HIV-1 and 15.0 years (CI 13.3-16.7) in HIV-2 infected individuals ($p<0.0001$, Log-Rank test). These analyses strongly support and demonstrate the robustness of our original estimates of survival time and progression time to AIDS in HIV-1 and HIV-2 infection.

2. Analyses of rate of follow-up

The number of individuals with at least one follow-up visit estimates the participation frequency in the study after initial enrollment (Table S1). However, it is also important to estimate the rate with which participants attended the

offered follow-up visits over the time under observation. Inspection of Q-Q-plots showed that the rates of follow-up (control visits/year) were not normally distributed, suggesting that non-parametric statistics should be used for this parameter. The median number of control visits/year in the entire cohort was 0.63 (IQR 0.47-0.87) for individuals with at least one follow-up visit (n=3761).

The median number of control visits/year for the individuals included in the current study where an estimated seroconversion date was available was 0.54 (IQR 0.42-0.65) for HIV-1 infected individuals (n=225) and 0.48 (IQR 0.36-0.60) for HIV-2 infected individuals (n=87). This analysis indicated similar follow-up rates (number of control visits/year) between HIV-1 and HIV-2 infected individuals.

3. Analyses of loss to follow-up

Loss to follow-up loosely means that the event status of a subject cannot be determined. However, the operational definition that is relevant for a specific study is a more difficult question and there is no clear consensus. Our study ended February, 2011. However, follow-up of the HIV-1 and HIV-2 positive individuals infected after enrollment continued until September 2013. Because follow-up visits were offered at a time interval of 12-18 months, the actual last opportunity for follow-up varied between subjects. Also, study participants that missed one opportunity for a control visit often returned at the next opportunity. Taking this into account we may define loss to follow-up to include all event-free individuals that did not attend a follow-up visit within the last 3 years (2 x 18 months). For comparison, 90% (3402 of 3761 participants) of the individuals in the cohort with at least one follow-up have a mean time interval between visits of three years or less. Thus, 90% of the study participants would be expected to attend a control visit within the last 3 years. Considering the median follow-up rate of 0.63 and the long follow-up period of the cohort (>20 years), we believe that this definition of lost to follow-up is reasonable. In addition, it would be similar to the "time since last clinical visit" or "lateness" definition that is often used, for example, in ART studies⁹. According to this definition, the number of individuals lost to follow-up (LTFU) in the cohort would be 2145 out of 4817 (45%). Among the individuals with at least one follow-up visit the LTFU was 1374 of 3761 (37%). LTFU stratified by inclusion date is presented in Table S7.

Table S7. Number of individuals lost to follow-up for the entire cohort. Number of individuals enrolled and with at least one follow-up (FU) visit (percentage of enrolled in parenthesis). Loss to follow-up is here defined as the number of individuals censored at a time before 3 years before the study end (LTFU). The cohort is stratified by inclusion date.

Inclusion period	Enrolled	≥ 1 FU visit (%)	LTFU (enrolled) (%)	LTFU (≥ 1 FU visit) (%)
1990-1994	1984	1798 (90%)	993 (50%)	807 (45%)
1995-1999	820	670 (82%)	494 (60%)	345 (51%)
2000-2004	1234	936 (76%)	490 (40%)	197 (21%)
2005-2009	779	357 (46%)	168 (22%)	25 (7%)
Total	4817	3761 (78%)	2145 (45%)	1374 (37%)

Next, we analysed the number of HIV seroincident individuals lost to follow-up stratified by inclusion date (Table S8). The analysis showed that there was no statistical difference in numbers lost to follow-up between HIV-1 and HIV-2 infected individuals, regardless of inclusion period.

Table S8. Number of individuals lost to follow-up for HIV seroincident individuals stratified by inclusion period. Number of HIV seroincident cases (N) and the number of individuals lost to follow-up defined as the number of individuals censored at a time before 3 years before the study end (LTFU). Data is stratified by inclusion date. P-values from two-tailed Fisher's exact test (FET) comparing the number of individuals LTFU between the two groups are presented.

Inclusion period	N		LTFU (%)		p (FET)
	HIV-1	HIV-2	HIV-1	HIV-2	
1990-1994	134	74	22 (16%)	18 (24%)	0.20

1995-1999	31	8	4 (13%)	3 (38%)	0.14
2000-2004	44	3	7 (16%)	0 (0%)	1.00
2005-2009	16	2	0 (0%)	0 (0%)	1.00
Total	225	87	33 (15%)	21 (24%)	0.07

Table S9 shows the numbers lost to follow-up stratified by the date of seroconversion. No statistical difference was found in the proportions lost to follow-up between HIV-1 and HIV-2 infected individuals in any of the analysed time periods.

Table S9. Number of individuals lost to follow-up for HIV seroincident individuals stratified by HIV seroconversion period. Number of HIV seroincident cases (N) and the number of individuals lost to follow-up defined as the number of individuals censored at a time before 3 years before the study end (LTFU). Data is stratified by HIV seroconversion period. P-values from two-tailed Fisher's exact test (FET) comparing the number of individuals LTFU between the two groups are presented.

HIV seroconversion period	N		LTFU (%)		p (FET)
	HIV-1	HIV-2	HIV-1	HIV-2	
1990-1999	90	61	19 (21%)	18 (30%)	0.25
2000-2009	135	26	14 (10%)	3 (12%)	0.74
Total	225	87	33 (15%)	21 (24%)	0.07

Another common approach, that we believe is most relevant to our study, is to estimate the LTFU over specific time intervals. The unique aspect of our study is the large number of seroincident cases (i.e. individuals with a seroconversion recorded after inclusion in the cohort) with long follow-up. Hence, these participants have, by definition, at least one follow-up visit. For comparison, it therefore makes sense to calculate LTFU in the total cohort for the individuals with at least one follow-up visit. The estimates of LTFU over specified time intervals since inclusion gives a picture of the general attrition rate in the cohort (Table S10). LTFU is generally low, except for the longest time intervals, considering that this is a prospective open cohort that has been followed for almost 20 years.

Table S10. Loss to follow-up over time intervals since inclusion for the entire cohort. Loss to follow-up (LTFU) was defined as the event-free individuals with at least one follow-up visit censored within 3, 5, 10 or >10 years since initial enrollment but at a time before 3 years before the study end (LTFU3, LTFU5, LTFU10 and LTFU>10, respectively). The LTFU values are also presented as percentages of the number of individuals with at least one follow-up visit in parenthesis.

Inclusion period	LTFU3 (%)	LTFU5 (%)	LTFU10 (%)	LTFU>10 (%)
1990-1994	163 (9%)	291 (16%)	490 (27%)	807 (45%)
1995-1999	130 (19%)	147 (22%)	311 (46%)	345 (51%)
2000-2004	190 (20%)	197 (21%)	197 (21%)	197 (21%)
2005-2009	25 (7%)	25 (7%)	25 (7%)	25 (7%)
Total	508 (14%)	660 (18%)	1023 (27%)	1374 (37%)

The LTFU over specific time intervals from seroconversion for the HIV seroincident individuals are shown in Table S11. The attrition rates for the seroincident individuals are generally slightly lower than in the entire cohort. There were no significant differences in LTFU between the HIV-1 and HIV-2 infected groups in any of the time intervals. Moreover, we consider the numbers lost to follow-up to be small (12% LTFU (38/312) within 5 years, 16% LTFU (51/312) within 10 years, and 17% LTFU (54/312) within >10 from seroconversion). These values should be compared to the median survival times of 8.2 years for HIV-1 and 15.6 years for HIV-2 infected individuals, and progression-times to AIDS of 6.2 years for HIV-1 and 14.3 years for HIV-2 infected individuals. It seems highly

unlikely that bias among the 12-17% of individuals that were lost to follow-up over the relevant time periods, would have flawed our results and conclusions.

Table S11. Loss to follow-up over time intervals since seroconversion for HIV seroincident individuals. Loss to follow-up (LTFU) was defined as the event-free individuals with at least one follow-up visit censored within 3, 5, 10 or >10 years since HIV seroconversion but at a time before 3 years before the study end (LTFU3, LTFU5, LTFU10 and LTFU>10, respectively). The LTFU values are also presented as percentages of the number of individuals with at least one follow-up visit in parenthesis. P-values from two-tailed Fisher's exact test (FET) comparing the number of individuals LTFU between the two groups are presented.

HIV seroconversion period		LTFU3 (%)	LTFU5 (%)	LTFU10 (%)	LTFU>10 (%)
1990-1999	HIV-1	4/90 (4%)	11/90 (12%)	19/90 (21%)	19/90 (21%)
	HIV-2	5/61 (8%)	10/61 (16%)	15/61 (25%)	18/61 (30%)
	p-value	p=0.49	p=0.48	p=0.69	p=0.25
2000-2009	HIV-1	10/135 (7%)	14/135 (10%)	14/135 (10%)	14/135 (10%)
	HIV-2	2/26 (8%)	3/26 (12%)	3/26 (12%)	3/26 (12%)
	p-value	p=1.0	p=0.74	p=0.74	p=0.74
Total	HIV-1	14/225 (6%)	25/225 (11%)	33/225 (15%)	33/225 (15%)
	HIV-2	7/87 (8%)	13/87 (15%)	18/87 (21%)	21/87 (24%)
	p-value	p=0.62	p=0.34	p=0.23	p=0.065

4. Detailed description of HIV seroincident individuals lost to follow-up

Next, we specifically examined the individuals that were considered to be lost to follow-up among the HIV seroincident individuals. One of the assumptions in survival analysis is that the characteristics of the groups under study and the predictors of loss to follow-up (other than the investigated factor) are equal.

First, there were no significant differences in time under observation between the individuals lost to follow-up in the HIV-1 and the HIV-2 groups and the distribution of individuals in the different intervals of time under observation was similar (Table S12).

Table S12. Intervals of time under observation from seroconversion for the seroincident individuals lost to follow-up. The numbers of HIV seroincident HIV-1 and HIV-2 infected individuals that were censored at a time before 3 years before the study end stratified by seroconversion date and time under observation (TUO). TUO was estimated as the time from inclusion to last follow-up visit. Two-tailed Fisher's exact test (FET) was used to test for any statistical differences between groups in each interval.

	HIV-1	HIV-2	P (FET)
>3 years of TUO	25/33 (76%)	18/21 (86%)	0.50
>5 years of TUO	23/33 (70%)	15/21 (71%)	1.00
>10 years of TUO	14/33 (42%)	9/21 (43%)	1.00
>15 years of TUO	1/33 (3%)	0/21 (0%)	1.00

We also investigated and found no indication of skew in age, sex or continuing participation rate (number of control visits/year) between the HIV-1 and HIV-2 groups (Table S13). In addition, the values (time under observation, number of control visits/year, age and sex) were similar to the corresponding values for all the seroincident individuals (Table S6, Table S7, and Table 1 in the main text).

Table S13. Characteristics of individuals that were censored earlier than 3 years before study closure (and thereby defined as lost to follow-up).

	HIV-1	HIV-2	Statistics
Mean age (SD¹)	37 (10)	39 (9)	p=0.64 ²
Median age (IQR³)	36 (30-44)	40 (31-45)	p=0.49 ⁴
Proportion Men (% men)	29/33 (88%)	17/21 (81%)	p=0.70 ⁵
Mean control visits/year since inclusion (SD¹)	0.80 (0.52)	0.67 (0.28)	p=0.25 ²
Median control visits/year since inclusion (IQR³)	0.60 (0.44-1.05)	0.58 (0.48-0.82)	p=0.75 ⁴

¹SD = Standard deviation

²Two-tailed student's T-test

³IQR = Interquartile range

⁴Two-tailed Mann-Whitney U test

⁵Two-tailed Fisher's exact test

Second, we repeated the survival analysis excluding the individuals lost to follow-up. The median survival times were 8.1 years for HIV-1 and 15.6 years for HIV-2 infected individuals ($p < 0.001$, Log-Rank test), and the median progression-times to AIDS were 6.2 years for HIV-1 and 13.3 years for HIV-2 infected individuals ($p < 0.001$, Log-Rank test). The estimates and differences between the groups were very similar to the original analyses, indicating that the impact of the individuals lost to follow-up was small.

5. Uncertainty in estimated HIV seroconversion date

In patient cohorts it is rare to have reliable data on the time of HIV infection, especially in West African cohorts. In our cohort, the time of infection was estimated to be the mid-point between the last HIV-negative and the first HIV-positive sample during follow-up. The time between the last negative and first positive sample can sometimes be long causing a variable and sometimes large uncertainty in the estimated infection date. If the true infection time point deviates systematically from the estimate, this could affect the estimates and the outcome of the comparison between the HIV-1 and HIV-2 infected groups if 1) the level of uncertainty in the estimate differs between groups or, 2) the degree of deviation of the estimate from the true time of infection differs between the groups. In our study, the average window of seroconversion were not statistically different between HIV-1 and HIV-2 infected individuals ($p=0.24$, 2-tailed student's T-test).

To control for this potential caveat of the survival estimates, we gradually removed cases with higher uncertainty in the estimated infection date. This was done both according to a Kaplan-Meier analysis and a Cox proportional regression analysis controlling for age and sex (Tables S14 and S15). Despite a large reduction of the original dataset, especially in the final steps, both analyses showed a consistent difference in survival time (Kaplan-Meier analysis) or Hazard ratio of mortality (Cox proportional regression analysis) between HIV-1 and HIV-2 infected individuals until the final step when only 4% and 6% of the original data sets remained. Only the last step showed a non-significant result (step 0-1 year uncertainty, $p=0.97$ for the Kaplan-Meier analysis and $p=0.84$ for the Cox proportional regression analysis). However, in this step the dataset is reduced to a level where it would be almost impossible to identify any statistical significant difference between the two groups (despite a large difference in effect size also at this level).

Table S14. Analysis of survival in reduced datasets with progressively smaller uncertainty in estimated HIV seroconversion date (Kaplan-Meier analysis).

Uncertainty in SC date ¹	Status	N	Percentage of the original dataset	Median survival time	Log-Rank ²
0-14 years (all samples)	HIV-1	225	100%	8.2 years	<0.001
	HIV-2	87	100%	15.6 years	
0-13 years	HIV-1	224	100%	8.2 years	<0.001
	HIV-2	87	100%	15.6 years	
0-12 years	HIV-1	220	98%	8.2 years	<0.001
	HIV-2	87	100%	15.6 years	
0-11 years	HIV-1	218	97%	8.2 years	<0.001
	HIV-2	85	98%	15.6 years	
0-10 years	HIV-1	218	97%	8.2 years	<0.001
	HIV-2	84	97%	15.6 years	
0-9 years	HIV-1	214	95%	8.2 years	<0.001
	HIV-2	83	95%	18.2 years	
0-8 years	HIV-1	211	94%	8.2 years	<0.001
	HIV-2	83	95%	18.2 years	
0-7 years	HIV-1	203	90%	8.1 years	<0.001
	HIV-2	79	91%	18.2 years	
0-6 years	HIV-1	194	86%	8.1 years	<0.001
	HIV-2	77	89%	18.2 years	
0-5 years	HIV-1	183	81%	8.1 years	<0.001
	HIV-2	69	79%	18.2 years	
0-4 years	HIV-1	167	74%	8.1 years	<0.001
	HIV-2	66	76%	18.2 years	
0-3 years	HIV-1	125	56%	8.1 years	<0.001
	HIV-2	57	66%	18.2 years	
0-2 years	HIV-1	78	35%	8.0 years	<0.001
	HIV-2	43	49%	18.2 years	
0-1 years	HIV-1	10	4%	10.4 years	0.97
	HIV-2	5	6%	8.0 years	

¹SC: Estimated seroconversion date. The uncertainty is referring to the elapsed time between the last HIV-negative and the first HIV-positive sample for each patient (that is: the longer time between those samples, the higher the uncertainty)

²p-value from Kaplan-Meier analysis (Log-Rank test)

Table S15. Analysis of Hazard ratio for mortality in reduced datasets with progressively smaller uncertainty in estimated HIV seroconversion date (Cox regression analysis controlling for age and sex).

Uncertainty in SC date ¹	Status	N	Percentage of the original dataset	Hazard ratio ²	Wald ³
0-14 years (all samples)	HIV-1	225	100%	3.50	<0.001
	HIV-2	87	100%		
0-13 years	HIV-1	224	100%	3.50	<0.001
	HIV-2	87	100%		
0-12 years	HIV-1	220	98%	3.53	<0.001
	HIV-2	87	100%		
0-11 years	HIV-1	218	97%	3.48	<0.001
	HIV-2	85	98%		
0-10 years	HIV-1	218	97%	3.55	<0.001
	HIV-2	84	97%		
0-9 years	HIV-1	214	95%	3.64	<0.001
	HIV-2	83	95%		
0-8 years	HIV-1	211	94%	3.69	<0.001
	HIV-2	83	95%		
0-7 years	HIV-1	203	90%	3.68	<0.001
	HIV-2	79	91%		
0-6 years	HIV-1	194	86%	3.36	<0.001
	HIV-2	77	89%		
0-5 years	HIV-1	183	81%	3.17	<0.001
	HIV-2	69	79%		
0-4 years	HIV-1	167	74%	3.08	<0.001
	HIV-2	66	76%		
0-3 years	HIV-1	125	56%	3.26	<0.001
	HIV-2	57	66%		
0-2 years	HIV-1	78	35%	3.32	<0.001
	HIV-2	43	49%		
0-1 years	HIV-1	10	4%	1.14	0.84
	HIV-2	5	6%		

¹SC: Estimated seroconversion date. The uncertainty is referring to the elapsed time between the last HIV-negative and the first HIV-positive sample for each patient (that is: the longer time between those samples, the higher the uncertainty)

²The adjusted hazard ratio of mortality in HIV-1 vs. HIV-2 infected individuals

³p-value from Cox regression analysis (Wald test)

The corresponding analyses for progression-time to AIDS showed similar consistencies as in the analysis of survival times (Tables S16 and S17) with a non-significant result seen only in the last step (step 0-1 year uncertainty, p=0.81 for the Kaplan-Meier analysis and p=0.31 for the Cox proportional regression analysis).

In summary, these results show that it is highly unlikely that the outcome of comparisons between the HIV-1 and HIV-2 infected groups were affected by any bias in the estimated seroconversion dates. At all steps, except for the last step of the reduction analysis, analysis of the remaining data did not deviate from the original results. Furthermore, there is no reason *a priori* to suspect that the true infection dates in either group deviated more or less from the estimated infection dates.

Table S16. Analysis of progression time to AIDS in reduced datasets with progressively smaller uncertainty in estimated HIV seroconversion date (Kaplan-Meier analysis).

Uncertainty in SC date ¹	Status	N	Percentage of the original dataset	Median progression-time to AIDS	Log-Rank ²
0-14 years (all samples)	HIV-1	225	100%	6.2 years	<0.001
	HIV-2	87	100%	14.3 years	
0-13 years	HIV-1	224	100%	6.2 years	<0.001
	HIV-2	87	100%	14.3 years	
0-12 years	HIV-1	220	98%	6.1 years	<0.001
	HIV-2	87	100%	14.3 years	
0-11 years	HIV-1	218	97%	6.1 years	<0.001
	HIV-2	85	98%	14.3 years	
0-10 years	HIV-1	218	97%	6.1 years	<0.001
	HIV-2	84	97%	14.3 years	
0-9 years	HIV-1	214	95%	6.1 years	<0.001
	HIV-2	83	95%	14.9 years	
0-8 years	HIV-1	211	94%	5.9 years	<0.001
	HIV-2	83	95%	14.9 years	
0-7 years	HIV-1	203	90%	5.7 years	<0.001
	HIV-2	79	91%	14.9 years	
0-6 years	HIV-1	194	86%	5.7 years	<0.001
	HIV-2	77	89%	14.9 years	
0-5 years	HIV-1	183	81%	5.4 years	<0.001
	HIV-2	69	79%	14.3 years	
0-4 years	HIV-1	167	74%	5.3 years	<0.001
	HIV-2	66	76%	15.0 years	
0-3 years	HIV-1	125	56%	5.3 years	<0.001
	HIV-2	57	66%	15.6 years	
0-2 years	HIV-1	78	35%	5.7 years	0.003
	HIV-2	43	49%	15.6 years	
0-1 years	HIV-1	10	4%	3.8 years	0.81
	HIV-2	5	6%	8.0 years	

¹SC: Estimated seroconversion date. The uncertainty is referring to the elapsed time between the last HIV-negative and the first HIV-positive sample for each patient (that is: the longer time between those samples, the higher the uncertainty)

²p-value from Kaplan-Meier analysis (Log-Rank test)

Table S17. Analysis of Hazard ratio of progression time to AIDS in reduced datasets with progressively smaller uncertainty in estimated HIV seroconversion date (Cox regression analysis controlling for age and sex).

Uncertainty in SC date ¹	Status	N	Percentage of the original dataset	Hazard ratio ²	Wald ³
0-14 years (all samples)	HIV-1	225	100%	2.84	<0.001
	HIV-2	87	100%		
0-13 years	HIV-1	224	100%	2.88	<0.001
	HIV-2	87	100%		
0-12 years	HIV-1	220	98%	2.94	<0.001
	HIV-2	87	100%		
0-11 years	HIV-1	218	97%	2.89	<0.001
	HIV-2	85	98%		
0-10 years	HIV-1	218	97%	2.89	<0.001
	HIV-2	84	97%		
0-9 years	HIV-1	214	95%	2.98	<0.001
	HIV-2	83	95%		
0-8 years	HIV-1	211	94%	3.02	<0.001
	HIV-2	83	95%		
0-7 years	HIV-1	203	90%	2.93	<0.001
	HIV-2	79	91%		
0-6 years	HIV-1	194	86%	2.80	<0.001
	HIV-2	77	89%		
0-5 years	HIV-1	183	81%	2.74	<0.001
	HIV-2	69	79%		
0-4 years	HIV-1	167	74%	3.07	<0.001
	HIV-2	66	76%		
0-3 years	HIV-1	125	56%	3.21	<0.001
	HIV-2	57	66%		
0-2 years	HIV-1	78	35%	2.67	0.001
	HIV-2	43	49%		
0-1 years	HIV-1	10	4%	2.44	0.31
	HIV-2	5	6%		

¹SC: Estimated seroconversion date. The uncertainty is referring to the elapsed time between the last HIV-negative and the first HIV-positive sample for each patient (that is: the longer time between those samples, the higher the uncertainty)

²The adjusted hazard ratio of progression to AIDS in HIV-1 vs. HIV-2 infected individuals

³p-value from Cox regression analysis (Wald test)

6. HIV-related mortality and time to WHO 4 stage

In this cohort, data on HIV-related death was also collected (6% of HIV-1 and 9% of HIV-2 single infected individuals with an estimated date of infection died of non-HIV-related causes, and were right censored at their last follow-up date). The median survival time was 8.8 years (95% confidence interval [CI] 7.9-9.8) in HIV-1 and 18.2 years (CI 14.3-22.1) in HIV-2 infected individuals using HIV-related mortality instead of all-cause mortality ($p < 0.0001$, Log-Rank test). The Hazard Ratio (HR) for mortality among HIV-1 compared with HIV-2 infected individuals was 3.64 in a Cox proportional-hazards model controlling for age and sex (CI 2.22-5.97, $p < 0.0001$, Wald test). Age (HR 1.03 per year of age increase, CI 1.00-1.05, $p = 0.024$, Wald test) and sex (HR 1.42 for men vs. women, CI 1.04-1.92, $p = 0.025$, Wald test) were independently associated with mortality.

The comparison between the HIV-1 and HIV-2 infected groups based solely on WHO 4 staging showed a similar effect size with a 2.1 times faster development of WHO 4 among HIV-1 compared with HIV-2 infected individuals (to be compared with 2.3 times faster development of AIDS in our original analysis taking the full scope of the WHO case definition into account). As expected, the estimates were slightly higher when only the WHO 4 staging was used: HIV-1: 8.7 years (CI 7.9-9.5) vs. HIV-2: 18.2 years (CI 13.2-23.2), $p < 0.0001$ (Log Rank test).

7. Potential bias of antiretroviral treatment initiation and HIV-1 and HIV-2 dual-infection

Further sensitivity analyses of potential bias induced by the small proportion of individuals starting antiretroviral treatment during follow-up were done by reanalysis of the dataset excluding the individuals that later initiated antiretroviral treatment. The analyses followed the same set-up of Kaplan-Meier and Cox regression analysis as presented in the analyses of the complete dataset. Seven HIV-1 and four HIV-2 infected individuals infected after enrollment initiated antiretroviral treatment during follow-up. The median survival time was 8.2 years (95% confidence interval [CI] 7.5-8.8) in HIV-1 and 15.6 years (CI 12.0-19.3) in HIV-2 infected individuals ($p < 0.0001$, Log-Rank test). The Hazard Ratio (HR) for mortality among HIV-1 compared with HIV-2 infected individuals was 3.33 in a Cox proportional-hazards model controlling for age and sex (CI 2.11-5.26, $p < 0.0001$, Wald test). The median time to AIDS was 6.4 years (95% CI 5.5-7.4) in HIV-1 and 14.9 years (CI 10.5-19.3) in HIV-2 infected individuals ($p < 0.0001$, Log-Rank test). The HR for time to AIDS among HIV-1 compared with HIV-2 infected individuals was 2.96 in a Cox proportional-hazards model controlling for age and sex (CI 1.95-4.48, $p < 0.0001$, Wald test).

Finally, we also performed a sensitivity analyses of potential bias induced by the small proportion of individuals becoming HIV-1 and HIV-2 dual-infected during follow-up by reanalysis of the dataset excluding the individuals that later became dual-infected. The analyses followed the same set-up of Kaplan-Meier and Cox regression analyses of the complete dataset. One HIV-1 and four HIV-2 infected individuals infected after enrollment became dual-infected during follow-up. The median survival time for was 8.2 years (CI 7.5-8.9) in HIV-1 and 15.6 years (CI 12.0-19.2) in HIV-2 infected individuals ($p < 0.0001$, Log-Rank test). The Hazard Ratio (HR) for mortality among HIV-1 compared with HIV-2 infected individuals was 3.50 in a Cox proportional-hazards model controlling for age and sex (CI 2.21-5.52, $p < 0.0001$, Wald test). The median time to AIDS was 6.3 years (95% CI 5.4-7.2) in HIV-1 and 14.3 years (CI 9.7-19.0) in HIV-2 infected individuals ($p < 0.0001$, Log-Rank test). The HR for time to AIDS among HIV-1 compared with HIV-2 infected individuals was 2.82 in a Cox proportional-hazards model controlling for age and sex (CI 1.89-4.20, $p < 0.0001$, Wald test).

In summary, these analyses clearly show the marginal impact of censoring due to antiretroviral therapy or dual-infection on the estimates of survival and times to AIDS.

8. Survival analysis stratified by sex

No significant differences in median follow-up time were found between men and women among either HIV-1 (8.8 years for men [CI 3.7-13.5] and 7.7 years for women [CI 4.1-10.7], $p = 0.52$, Log-Rank test) or HIV-2 infected individuals (10.4 years for men [CI 4.9-15.3] and 11.0 years for women [CI 3.7-16.2], $p = 0.92$, Log-Rank test). Men were older than women at estimated HIV-infection date in both HIV-infected groups (median age was 37.4 years among HIV-1 infected men, 32.6 years among HIV-1 infected women [$p = 0.0073$, two-tailed Mann-Whitney U test], 39.2 years among HIV-2 infected men, and 33.4 years among HIV-2 infected women [$p = 0.022$]).

A stratified analysis showed that the median survival time was 8.0 years for HIV-1 infected men (CI 7.4-8.5) compared to 11.3 years for HIV-1 infected women (CI 5.8-16.9, $p = 0.019$, Log-Rank test), and 15.0 years for HIV-2 infected men (CI 12.2-17.8) compared with 18.8 years for HIV-2 infected women (CI 12.4-25.3, $p = 0.064$, Log-Rank test). A Cox proportional-hazards model of HIV-1 infected individuals with age and sex as covariates showed that age (HR 1.03 per year of age increase, CI 1.01-1.06, $p = 0.0045$, Wald test) but not sex (HR 1.86 for men vs. women, CI 0.99-3.52, $p = 0.055$, Wald test) were significantly associated with mortality. The corresponding analysis for HIV-2 infected individuals showed no significant associations between age (HR 1.02 per year of age increase, CI 0.98-1.06, $p = 0.38$, Wald test) or sex (HR 3.40 for men vs. women, CI 0.79-14.7, $p = 0.10$, Wald test) with mortality. No significant interactions were found between age and sex.

The analysis stratified by sex showed that the median time to AIDS was 6.1 years for HIV-1 infected men (CI 5.1-7.2) compared with 8.2 years for HIV-1 infected women (CI 6.4-10.0, $p = 0.12$, Log-Rank test), and 11.0 years for

HIV-2 infected men (CI 9.4-12.7) compared with 18.8 years for HIV-2 infected women (CI 12.4-25.3, $p=0.0091$, Log-Rank test). A Cox proportional-hazards model of HIV-1 infected individuals with age and sex as covariates showed that age (HR 1.03 per year of age increase, CI 1.01-1.05, $p=0.0062$, Wald test) but not sex (HR 1.36 for men vs. women, CI 0.81-2.30, $p=0.25$, Wald test) was significantly associated with time to AIDS. The corresponding analysis for HIV-2 infected individuals showed that sex (HR 5.10 for men vs. women, CI 1.22-21.74, $p=0.027$, Wald test) but not age (HR 1.02 per year of age increase, CI 0.99-1.06, $p=0.22$, Wald test) was significantly associated with time to AIDS. No significant interactions were found between age and sex.

9. Characteristics of participants HIV-infected at enrollment

During the study period, 183 individuals were HIV-1 infected and 377 individuals were HIV-2 infected at enrollment (Table S18).

Table S18. Characteristics of the study groups that were HIV-infected at enrollment¹.

	HIV-1	HIV-2
Number of individuals	183	377
Number of men (%)	139 (76%)	295 (78%)
Median age at enrollment (years, IQR)²	33 (27-42)	39 (33-46)
Median follow-up time (years, IQR)³	2.8 (0-5.8)	6.7 (1.1-17.5)
Recorded deaths (%)	42 (23%)	117 (31%)
Median time to death or censoring (years, IQR)⁴	2.4 (0.0-5.4)	4.7 (1.0-11.0)
Median age at death (years, IQR)⁵	38 (31-47)	49 (41-55)
Number of individuals developing AIDS (%)	74 (40%)	131 (35%)
Median time to AIDS or censoring (years, IQR)⁶	1.4 (0.0-4.7)	3.5 (0.8-9.3)

¹There were no trends or significant pairwise differences between groups except as noted (significance cut-off: $p<0.05$). IQR: Interquartile range

²HIV-1 vs. HIV-2: $p<0.0001$, two-tailed Mann-Whitney U test (M-W).

³HIV-1 vs. HIV-2: $p<0.0001$, Log-Rank test as estimated by reverse Kaplan-Meier estimator⁷.

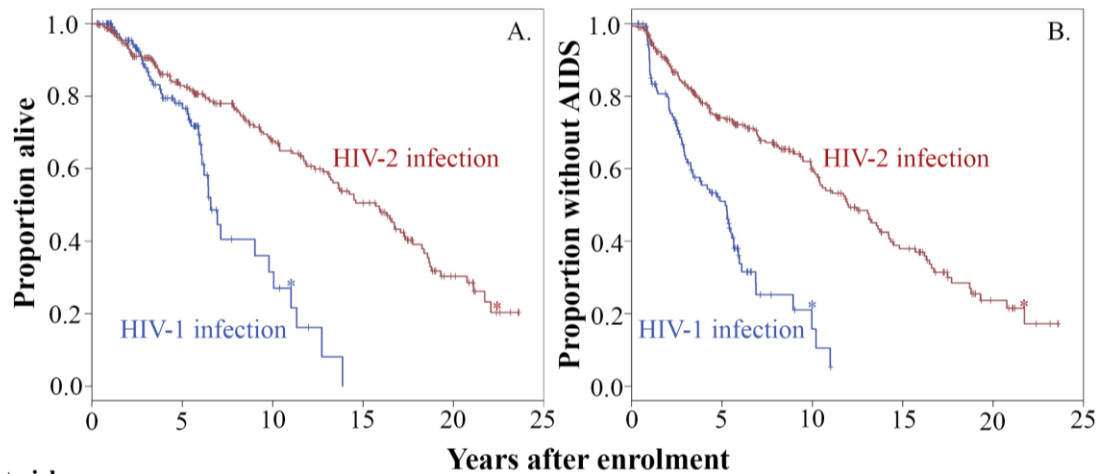
⁴HIV-1 vs. HIV-2: $p<0.0001$, M-W.

⁵HIV-1 vs. HIV-2: $p<0.0001$, M-W.

⁶HIV-1 vs. HIV-2: $p<0.0001$, M-W.

The median survival time from enrollment was 6.6 years (CI 5.8-7.4) in HIV-1 and 15.7 years (CI 13.5-17.9) in HIV-2 infected individuals ($p<0.0001$, Log-Rank test, Figure S2A). The hazard ratio for mortality among HIV-1 compared with HIV-2 infected individuals was 2.75 in a Cox proportional-hazards model controlling for age and sex (CI 1.86-4.07, $p<0.0001$, Wald test). Age (HR 1.03 per year of age increase, CI 1.01-1.05, $p=0.0085$, Wald test) but not sex (HR 1.05 for men vs. women, CI 0.85-1.31, $p=0.63$, Wald test) was independently associated with mortality.

The median time from enrollment to AIDS was 5.2 years (95% CI 3.9-6.5) in HIV-1 and 11.2 years (CI 9.9-14.1) in HIV-2 infected individuals ($p<0.0001$, Log-Rank test, Figure S2B). The HR for time to AIDS among HIV-1 compared with HIV-2 infected individuals was 3.19 in a Cox proportional-hazards model controlling for age and sex (CI 2.30-4.44, $p<0.0001$, Wald test). Neither age (HR 1.01 per year of age increase, CI 0.99-1.02, $p=0.38$, Wald test) nor sex (HR 1.09 for men vs. women, CI 0.91-1.31, $p=0.34$, Wald test) were significantly associated with time to AIDS.



Number at risk											
HIV-1 infection	183	53	7	0	0	0	183	42	1	0	0
HIV-2 infection	377	180	102	62	18	0	377	156	83	41	11

Figure S2. Kaplan-Meier curves of survival (A) and AIDS-free time (B) of individuals infected with HIV-1 or HIV-2 at enrollment. Tick marks indicate participants with censored data. Asterisks indicate the time-point in each group when five participants are still at mortality risk and risk of developing AIDS. The number of individuals at risk are given below the figures at 5-year intervals.

The incubation time from infection to development of AIDS can last for several years and lack of a reasonable estimate of time of infection can severely bias studies of natural HIV disease progression. Hence, individuals with a fast disease progression due to the development of AIDS or death relatively early after infection may be underrepresented among participants that are HIV-infected at enrollment (i.e. the likelihood of inclusion in the cohort increases with disease-free time). It is therefore possible that the participants that were HIV-infected at enrollment, at least in part, represent a selected group of disease controllers.

10. Trajectory of time to AIDS in HIV-1 and HIV-2 infection

Fitting Weibull models for time to AIDS indicated that the curves mainly differed by the scale parameters (8.23 [CI 7.42-9.12] for HIV-1 and 15.95 [CI 11.24-22.64] for HIV-2), and that the shape parameters were almost identical (HIV-1: 1.73 [CI 1.52-1.97]; HIV-2: 1.59 [CI 1.10-2.11], Figure S3A). Model comparison by Aikike Information Criteria (AIC) suggested that a common shape parameter on separate time scales provided the best fit model to describe the time to AIDS functions of HIV-1 and HIV-2 infected individuals (AIC=1018.7, compared with AIC=1051.3 for separate shape, same scale; and AIC=1019.8 for separate shape, separate scale). The density plots peaked marginally before the estimated median survival time with a slightly right-tailed distribution (Figure S3B). To further explore the similarity in shape of the Weibull distributions, we estimated the acceleration factor θ to 1.94 (CI 1.51-2.48). To ease the visual comparison of survival curves between HIV-1 and HIV-2 infected individuals, the times to AIDS of the HIV-1 infected individuals were multiplied by θ (Figure S3C). Almost identical shapes and overlapping confidence intervals throughout the disease course suggest high similarities between HIV-1 and HIV-2 infection.

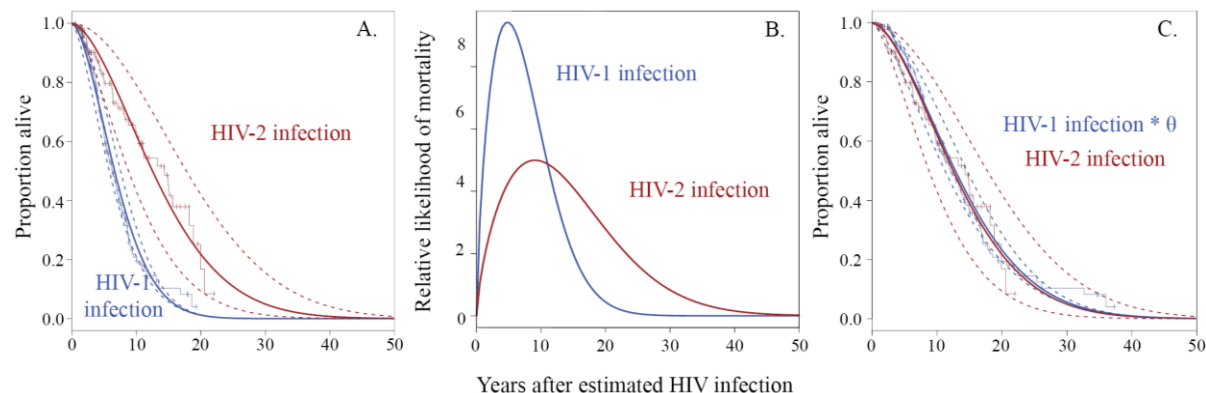


Figure S3. Characteristics of time to AIDS in HIV-1 and HIV-2 infection. (A) Weibull estimates of the time to AIDS function (solid lines) with 95% confidence intervals (dashed lines) for HIV-1 (blue) and HIV-2 (red) infected individuals. Kaplan-Meier curves are plotted for comparison. (B) Probability density functions of the Weibull distributions showing the relative likelihood of developing AIDS in the infected population over time. (C) Kaplan-Meier and Weibull curves (with 95% confidence intervals, dashed lines) for HIV-2 infected individuals (red) compared to a simulated HIV-2 dataset, generated by multiplying the times to AIDS of the HIV-1 infected individuals by the acceleration factor θ determined by comparing HIV-1 and HIV-2 infected individuals in a Weibull regression model (blue).

11. Analysis of T-cell dynamics

It has been suggested that percentages of T-cell populations may be more suitable than absolute CD4⁺ T-cell counts in resource-limited settings because of their lower variability and lower sensitivity to specimen handling, participant age, and time of sampling¹⁰⁻¹². Hence, the analyses of T-cell dynamics in the main text focus on percentages. However, corresponding analyses were also done for absolute T-cell counts and are presented below.

The T-cell populations in untreated HIV-infection are generally considered to change in a linear fashion over time¹³. However, this assumption is based on HIV-1 data and to the best of our knowledge, alternative models have not been thoroughly tested for HIV-2 infection. We therefore applied ten different models (linear, logarithmic, inverse, quadratic, cubic, compound, power, S, growth, exponential, and logistic) available in IBM® SPSS® Statistics version 23 (IBM Corporation) on time series of CD4% data from HIV-1 and HIV-2 infected individuals, respectively. Assessment by Pearson's R and residuals indicated similar goodness of fit of the models and all models gave reasonable intercepts of CD4%. The analysis showed that none of the evaluated models had a statistically better fit for either the HIV-1 or the HIV-2 data compared with the linear model. Although the linear model was not significantly better than the majority of the tested models either, we decided to use a full-factorial linear mixed model with time since infection as covariate for comparability with previous literature¹¹.

Analysis of CD4% by sex showed similar rates of decline in men and women for both HIV-1 and HIV-2 infected individuals, respectively (Table S19). However, the mean level of CD4% at the estimated date of infection was significantly higher among HIV-2 infected women (35.0%) compared with HIV-2 infected men (26.7%, $p=0.022$, LRT). This sex difference was less pronounced among HIV-1 infected individuals ($p=0.057$, LRT). No significant differences in CD8% changes over time or levels were found when stratified by sex (Table S1).

Table S19. CD4%¹ and CD8%¹ for HIV-1 and HIV-2 infected individuals².

	HIV-1		HIV-2		HIV-1	HIV-1	HIV-2	HIV-2	
	Mean	Mean	P value	Mean	Mean	P value	Mean	Mean	P value
	(SE)	(SE)		(SE)	(SE)		(SE)	(SE)	
CD4% decline over time (per year)	0.9 (0.2)	0.4 (0.2)	0.025	0.9 (0.2)	0.8 (0.4)	0.83	0.3 (0.2)	0.4 (0.4)	0.72
CD4% at intercept	22.3 (1.7)	28.0 (1.3)	0.00094	21.4 (1.2)	25.8 (2.3)	0.057	26.7 (1.4)	35.0 (3.6)	0.022
CD8% increase over time (per year)	2.7 (0.5)	0.7 (0.3)	0.00015	2.7 (0.4)	2.3 (0.7)	0.55	1.1 (0.5)	-0.1 (1.1)	0.28
CD8% at intercept	38.3 (2.7)	34.9 (2.0)	0.21	38.3 (2.2)	38.4 (4.4)	0.99	33.8 (2.2)	35.6 (4.9)	0.72

¹CD4% and CD8% is defined as the CD4⁺ or CD8⁺ T-cell percentage of total lymphocytes.

²Two or more measurements of CD4⁺ T-cell percentage and CD8⁺ T-cell percentage were available for 85 (19 women, 22%) and 60 (10 women, 17%) HIV-1 infected individuals (in total 267 and 164 measurements), and 48 (7 women, 15%) and 39 (7 women, 18%) HIV-2 infected individuals (in total 186 and 125 measurements), respectively.

During follow-up, T-cell counts were also measured among approximately 10% of the HIV-negative population. We used this data as a reference of T cell levels among healthy individuals in the same cohort (Table S20). For the few individuals with more than one available T-cell count available, the first measurement was used in the analysis. The results suggested higher mean CD4⁺ T-cell levels (indicating better immune status) and lower CD8⁺ T-cell levels (indicating a lower degree of general immune activation) among HIV-negative individuals.

Table S20. CD4 and CD8 T-cell counts for HIV-negative individuals.

	Number of individuals tested	Mean (SE)
CD4%	313	38.0 (0.5)
Absolute CD4⁺ T-cell count	294	879.7 (23.4)
CD8%	304	30.3 (0.6)
Absolute CD8⁺ T-cell count	285	723.3 (27.5)

The mean rate of decline in absolute CD4⁺ T-cell counts was 22.5 per year for HIV-1 and 12.8 per year for HIV-2 infected individuals (p=0.060, LRT, Table S21). The mean absolute CD4⁺ T-cell levels at estimated date of infection were 447.9 for HIV-1 and 570.5 for HIV-2 infected individuals (p=0.012, LRT). No significant differences were found in absolute CD8⁺ T-cell counts between HIV-1 and HIV-2 infected individuals (Table S21). Stratification by sex showed no significant differences in change over time or in levels at estimated date of infection of absolute CD4⁺ T-cell counts among HIV-1 infected individuals (Table S21). In contrast, HIV-2 infected men and women had significant differences in both rate of decline (6.8 CD4⁺ T-cells per year for men and 35.3 CD4⁺ T-cells per year for women, p=0.0020, LRT) and intercept (495.8 CD4⁺ T-cells per µl blood for men and 929.7 CD4⁺ T-cells per µl blood for women, p=0.00031, LRT). No significant differences in absolute CD8⁺ T-cell changes over time or levels were found when stratified by sex (Table S21).

Table S21. Absolute CD4⁺ and CD8⁺ T-cell counts¹ for HIV-1 and HIV-2 infected individuals².

	HIV-1		HIV-2		HIV-1 Men	HIV-1 Women	HIV-2 Men		HIV-2 Women
	Mean (SE)	Mean (SE)	p value	Mean (SE)	Mean (SE)	p value	Mean (SE)	Mean (SE)	P value
Absolute CD4⁺ T-cell decline over time (per year)	22.5 (5.1)	12.8 (3.2)	0.060	22.6 (4.2)	27.2 (6.8)	0.50	6.8 (4.2)	35.3 (9.1)	0.0020
Absolute CD4⁺ T-cell count at intercept	447.9 (48.5)	570.5 (36.8)	0.012	443.4 (31.7)	482.5 (59.5)	0.51	495.8 (46.7)	929.7 (115.2)	0.00031
Absolute CD8⁺ T-cell increase over time (per year)	31.0 (27.3)	1.8 (18.7)	0.29	33.6 (27.6)	12.4 (53.4)	0.69	8.1 (14.6)	-15.4 (35.2)	0.51
Absolute CD8⁺ T-cell count at intercept	839.9 (128.5)	715.4 (91.1)	0.33	867.2 (113.7)	749.2 (212.9)	0.58	670.3 (88.5)	864.1 (203.2)	0.34

¹Absolute CD4⁺ and CD8⁺ T-cell counts in cells per µl blood.

²Two or more measurements of absolute CD4⁺ and CD8⁺ T-cell counts were available for 83 (19 women, 23%) and 58 (10 women, 17%) HIV-1 infected individuals (in total 258 and 155 measurements), and 47 (7 women, 15%) and 39 (7 women, 18%) HIV-2 infected individuals (in total 181 and 122 measurements), respectively.

To address any potential bias in our estimates of T-cell dynamics in HIV-1 and HIV-2 infection, we performed several control analyses. First, we analysed the complete dataset (i.e. including also individuals with only one measurement of CD4% during follow-up). This analyses included 140 HIV-1 and 72 HIV-2 infected individuals with an estimated HIV infection date (322 and 210 CD4% measurements, respectively). The analysis showed almost identical estimates as seen in the analysis of individuals with two or more measurements of CD4%. The mean rate of decline in CD4% was 0.9% (Standard error [SE] ±0.2%) for HIV-1 and 0.3% (SE ±0.1%) per year for HIV-2 infected individuals (p=0.020, likelihood-ratio test [LRT]). The extrapolated mean CD4% level at estimated date of infection was 21.8% (SE ±1.5%) for HIV-1 and 28.0% (SE ±1.1%) for HIV-2 infected individuals (p<0.0001, LRT). Second, to address how well the individuals with multiple measurements of CD4% represented the complete dataset, we performed a Kaplan-Meier analysis for survival and time to AIDS among the individuals with two or more measurements of CD4%. The median survival time was 10.6 years (95% confidence interval [CI] 8.7-12.6) in HIV-1 and 21.1 years (CI 13.3-28.9) in HIV-2 infected individuals (p<0.0001, Log-Rank test). The median time to AIDS was 6.3 years (95% CI 4.7-7.9) in HIV-1 and 13.3 years (CI 7.6-19.0) in HIV-2 infected individuals (p<0.0001, Log-Rank test). The estimates were slightly higher compared with the estimates from the analyses of the entire dataset. This was expected considering that individuals with longer follow-up are more likely to have two or more measurements of CD4%. All comparisons of survival or time to AIDS between the entire HIV-1 infected group with estimated infection date and the HIV-1 infected individuals with estimated infection date and two or more measurements of CD4%, and the corresponding analysis for HIV-2, showed non-significant results (in total four comparisons). The only exception was the comparison of survival between the entire HIV-1 infected group with estimated infection date and the HIV-1 infected individuals with estimated infection date and two or more measurements of CD4%. Altogether, these control analyses suggest that our estimates of T-cell dynamics in HIV-1 and HIV-2 infection are likely to be representative of the entire HIV-1 and HIV-2 infected cohort population, respectively.

12. CD4⁺ T-cell levels at AIDS

The median CD4% at clinical AIDS was 8.1% (IQR 3.1-12.4) for individuals infected with HIV-1 after enrollment and 8.3% (IQR 2.5-17.0) for individuals infected with HIV-1 before enrollment (p=0.69, two-tailed Mann-Whitney U test, Figure S4A). The corresponding numbers for HIV-2 were 25.8% (IQR 12.7-29.5) for individuals infected after enrollment and 17.2% (IQR 7.1-24.0) for individuals infected before enrollment (p=0.15, two-tailed Mann-Whitney U test, Figure S4A). The median absolute CD4⁺ T-cell level at clinical AIDS was 113.6 (IQR 33.8-284.1) for individuals infected with HIV-1 after enrollment and 145.1 (IQR 60.5-262.7) for individuals infected with HIV-1

before enrollment ($p=0.56$, two-tailed Mann-Whitney U test, Figure S4B). The corresponding numbers for HIV-2 were 215.6 (IQR 89.8-766.1) for individuals infected after enrollment and 251.3 (IQR 54.4-634.7) for individuals infected before enrollment ($p=0.85$, two-tailed Mann-Whitney U test, Figure S4B).

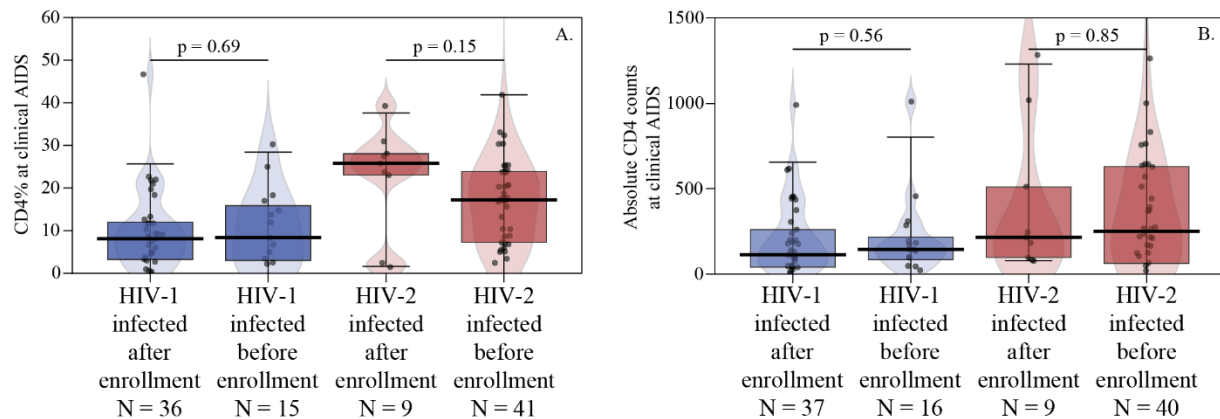


Figure S4. Pirate plots of estimated CD4% levels (A) and absolute CD4⁺ T-cell counts (B) at clinical AIDS for HIV-1 (blue) and HIV-2 (red) infected individuals stratified by HIV status at enrollment. Clinical AIDS was determined according to the World Health Organization clinical staging system¹⁴. Pirate plots combining individual observations (black points), box plots of the median, interquartile range and 95% quantile range, and descriptive smoothed density distributions of the data are shown for each group. Comparison between groups were performed by the two-tailed Mann-Whitney U Test.

The median CD4% at clinical AIDS was 8.2% (IQR 3.0-13.8) for HIV-1 and 18.2% (IQR 7.2-25.4) for HIV-2 infected individuals ($p<0.0001$, two-tailed Mann-Whitney U test, Figure 3 and S5). Comparison between men and women, showed that the median CD4% at clinical AIDS was 8.3% (IQR 2.7-14.0) for HIV-1 infected men and 8.0% (IQR 4.2-15.8) for HIV-1 infected women ($p=0.76$, two-tailed Mann-Whitney U test, Figure S5). The corresponding numbers for HIV-2 were 20.3% (IQR 6.9-25.6) for men and 17.2% (IQR 12.9-19.5) for women ($p=0.66$, two-tailed Mann-Whitney U test, Figure S5).

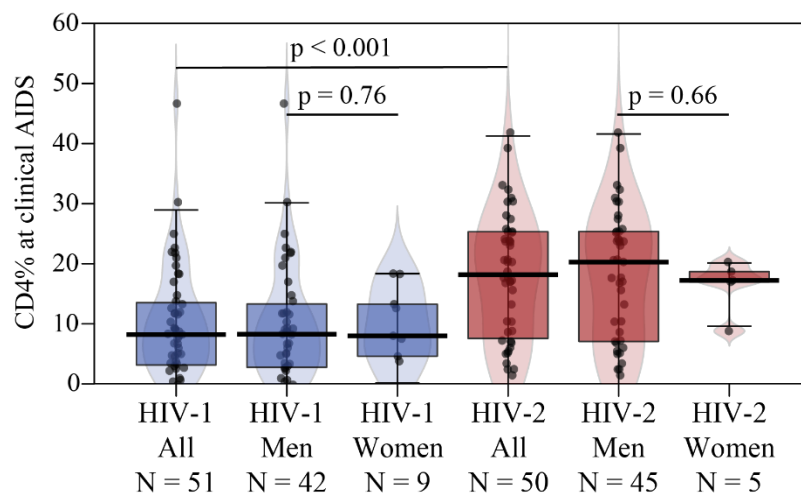


Figure S5. Pirate plots of estimated CD4% levels at clinical AIDS for HIV-1 and HIV-2 infected individuals stratified by sex. Clinical AIDS was determined according to the World Health Organization clinical staging system¹⁴. Both HIV-1 (blue) and HIV-2 infected individuals (red) were stratified and analysed by sex. Pirate plots combining individual observations (black points), box plots of the median, interquartile range and 95% quantile

range, and descriptive smoothed density distributions of the data are shown for each group. Comparison between groups were performed by the two-tailed Mann-Whitney U Test.

The median absolute CD4⁺ T-cell level at clinical AIDS was 136.8 (IQR 44.9-273.8) for HIV-1 and 236.7 (IQR 80.1-632.8) for HIV-2 infected individuals ($p=0.017$, two-tailed Mann-Whitney U test, Figure S6). Comparison between men and women, showed that the median absolute CD4⁺ T-cell level at clinical AIDS was 136.8 (IQR 44.8-306.5) for HIV-1 infected men and 126.4 (IQR 37.1-204.1) for HIV-1 infected women ($p=0.76$, two-tailed Mann-Whitney U test, Figure S6). The corresponding numbers for HIV-2 were 236.7 (IQR 77.7-636.6) for men and 307.0 (IQR 117.0-576.1) for women ($p=0.82$, two-tailed Mann-Whitney U test, Figure S6).

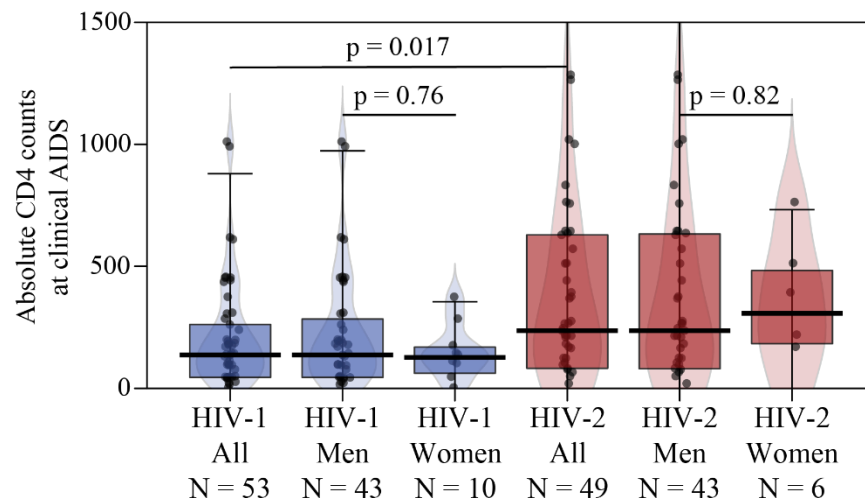


Figure S6. Pirate plots of estimated absolute CD4⁺ T-cell counts at clinical AIDS for HIV-1 (blue) and HIV-2 (red) infected individuals stratified by sex. Clinical AIDS was determined according to the World Health Organization clinical staging system¹⁴. Both HIV-1 (blue) and HIV-2 infected individuals (red) were stratified and analysed by sex. Pirate plots combining individual observations (black points), box plots of the median, interquartile range and 95% quantile range, and descriptive smoothed density distributions of the data are shown for each group. Comparison between groups were performed by the two-tailed Mann-Whitney U Test.

Supplementary References

1. Albert J, Bredberg U, Chiodi F, et al. A new human retrovirus isolate of West African origin (SBL-6669) and its relationship to HTLV-IV, LAV-II, and HTLV-IIIB. *AIDS Res Hum Retroviruses* 1987;3:3-10.
2. Andersson S, da Silva Z, Norrgren H, Dias F, Biberfeld G. Field evaluation of alternative testing strategies for diagnosis and differentiation of HIV-1 and HIV-2 infections in an HIV-1 and HIV-2-prevalent area. *AIDS* 1997;11:1815-22.
3. Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 1979;6:65-70.
4. RDCT. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2010.
5. Mansson F, Alves A, Silva ZJ, et al. Trends of HIV-1 and HIV-2 prevalence among pregnant women in Guinea-Bissau, West Africa: possible effect of the civil war 1998 1999. *Sex Transm Infect* 2007;83:463-7.
6. Mansson F, Biague A, da Silva ZJ, et al. Prevalence and incidence of HIV-1 and HIV-2 before, during and after a civil war in an occupational cohort in Guinea-Bissau, West Africa. *Aids* 2009;23:1575-82.
7. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996;17:343-6.
8. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part IV: further concepts and methods in survival analysis. *Br J Cancer* 2003;89:781-6.
9. Chi BH, Cantrell RA, Mwango A, et al. An empirical approach to defining loss to follow-up among patients enrolled in antiretroviral treatment programs. *Am J Epidemiol*;171:924-31.
10. Taylor JM, Fahey JL, Detels R, Giorgi JV. CD4 percentage, CD4 number, and CD4:CD8 ratio in HIV infection: which to choose and how to use. *J Acquir Immune Defic Syndr* 1989;2:114-24.
11. Esbjörnsson J, Månsson F, Kvist A, et al. Inhibition of HIV-1 disease progression by contemporaneous HIV-2 infection. *N Engl J Med* 2012;367:224-32.
12. Burcham J, Marmor M, Dubin N, et al. CD4% is the best predictor of development of AIDS in a cohort of HIV-infected homosexual men. *AIDS* 1991;5:365-72.
13. Margolick JB, Munoz A, Donnenberg AD, et al. Failure of T-cell homeostasis preceding AIDS in HIV-1 infection. The Multicenter AIDS Cohort Study. *Nature medicine* 1995;1:674-80.
14. WHO. WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children. Geneva: World Health Organization, 2007. <http://www.who.int/hiv/pub/guidelines/HIVstaging150307.pdf>.