Web Table 1: Review of all Published Studies Evaluating Cumulative HIV-Viremia as a Prognostic Predictor. As in the main text we use cVL_1 and cVL_2 to designate cumulative viral load metrics accumulated on a linear and log scale, respectively.

Study	Cumulative Viremia Calculation	Model[a]	Results[b]
Cole et al. 2009(3) MACS, USA, 1984-1998	**cVL_1:** Calculated by summing under the **viral load** curve (6 monthly measurements), **starting from seroconversion dates**, which were known and assumed to correspond to a viral load of zero.	Cox proportional hazards model of **progression to AIDS or mortality** (combined) vs. viral load measures including viral set point and time-updated values of log viral load, cVL_1 and peak viral load to date. Adjusted for time-updated CD4 (spline).	Baseline CD4[c] (at seroconversion): 701 (513-916) All four viral load measures significantly associated with hazard of AIDS/death in univariate models. Univariate model with the cVL_1 model chosen as the best univariate model by AIC selection. No viral load predictors significant in full multivariate model, perhaps partly due to collinearity. CD4 effect not shown.
Zoufaly et al. 2009(4) ClinSurv, Germany, 1999-2006	**cVL_2:** Calculated by summing under the **log viral load** curve (3 monthly measurements) and above the log (500copies/µl), **starting from ART initiation**; log (500) cutoff was chosen so that undetectable viral loads did not contribute to cumulative viremia. Baseline cumulative viremia assumed to be zero.	Cox proportional hazards model of **incident AIDS lymphoma** vs. time-updated viral load (categorized) and cVL_2 (continuous). Adjusted for baseline (< vs ≥ 200) and time-updated CD4 (< 200, 201-350, ≥350).	Baseline CD4 (at ART initiation) lymphoma: 90 (38-220) Baseline CD4 (at ART initiation) no lymphoma: 204 (80-340) Both time-updated viral load and cVL_2 significantly associated with hazard of AIDS lymphoma in univariate models. cVL_2 was also significantly associated with hazard in a multivariate model, but this model excluded time-updated viral load. Low baseline and time-updated CD4 also predictive of increased risk in multivariate model.
Marconi et al. 2011(5)	**cVL_1:** Calculated by summing under the **viral load** curve (6 monthly	Poisson regression of **progression to AIDS** vs. viral load decay rate over total	Baseline CD4 (at ART initiation): 278 (167-378)

Military HIV Natural History Study, USA 1986-2008	measurements) **starting from ART initiation**. cVL_1 was calculated only once per individual corresponding to their entire post-ART follow-up time (i.e. not time-updated).	follow-up time or within first year post-ART, viral load slope in first year post-ART and cVL_1. All four of these viral loads were dichotomized into binary variables based on their median values. Also included CD4 (continuous).	cVL_1, when included as a dichotomous but not continuous variable, was a statistically significant predictor of AIDS risk in either univariate or multivariate models. CD4 effect not shown.
Mugavero et al. 2011(6) CNICS, USA, 2000-2008	**cVL_1:** Calculated by summing under the **viral load** curve (6 monthly measurements) and then taking the logarithm, **starting from 24 weeks post-ART initiation**. In sensitivity analyses, summed cumulative viremia starting from ART start, 48 weeks and 2-years post-ART start.	Cox proportional hazards model of **all-cause mortality** vs. log viral load, cVL_1, log viral load at ART initiation and log viral load at 24 weeks post-ART initiation. Used marginal structural models to account for time-dependent confounding between viral load measures and CD4 counts.	Baseline CD4 (at ART initiation): 222 (97-325) Of four viral load measures in the multivariate adjusted model, only increasing cVL_1 was significantly associated with increased mortality risk. Lower time-updated CD4 was also associated with increased mortality risk. Sensitivity analyses were consistent (data not shown).
Mugavero et al. 2012(7) **UAB1917 clinic and UW Harborview Clinic, USA, 2007-2010**	**cVL_1;** Calculated by summing the area under the **viral load** curve (6 monthly measurements) and then taking the logarithm, starting from ART initiation and going up to two years.	Linear regression model of 2-year cVL_1 as an *outcome* variable as a function of clinic visit adherence, adjusting for baseline viral load and CD4, age, sex, race/ethnicity and health insurance.	Baseline CD4 at ART initiation: <200 (33%); 200-3500 (24%); and >350 (43%). Higher early retention rates were significantly associated with lower cVL_1 in a multivariate analysis.
Saracino et al. 2013(8)	Calculated by summing the area under the **viral**	Mann-Whitney test of cVL_1 over total time followed up as	Found significant differences in cVL_1 between patients infected with different strains.

Clinic of Infectious Diseases, Italy 1997-present	**load** curve (other details not provided).	an *outcome* variable as a function of HIV strain.	
Lima et al. 2014(9) **RCT NCT00162643, Mexico 2005-2007**	Calculated as total area under linear and log viral load curves (median (IQR) 5 (4-5) viral load test) each cVL measure included either baseline or ≥6months, to week 48.	Wilcoxon rank sum test analysis was used to assess association between cVL_1 and cVL_2 with viral load at 48 weeks post-ART initiation. Also assessed both cVL metrics as an *outcome* of randomized treatment assignment.	Median baseline CD4 count: median (IQR) 56cells/mm^3 (25-117) cVL_2 correlated with viral load at 48 weeks, though the former was derived from the latter. Patients initiated on efavirenz had significantly lower cVL_2 compared to lopinavir/r. cVL_1 did not significantly correlate with viral load at 48 weeks (likely because this measure is closely correlated with peak viral load, which generally occurs earlier post-ART initiation) or treatment assignment.
Kowalkowski et al. 2014(10) HIV-CCR, USA 1985-2010	**cVL_1;** Calculated by summing the area under the **viral load** curve (inconsistent inter-measurement duration, but averaging 3 per year) and then taking the logarithm, **starting from first observation**. Included individuals who had ever initiated treatment, including in the analysis person-time at risk pre-ART initiation.	Cox proportional hazards model for incidence of non-AIDS events (Hepatocarcinoma, Hodgkin lymphoma and squamous cell carcinoma of the anus) vs cVL_1, time-updated and pre-ART nadir CD4 count, log viral load and time-updated cumulative % of measurements with undetectable viral loads and several other variables.	Nadir CD4 (pre-ART initiation)[e]:<200 (45%); 200-350 (29%); and >350 (18%) cVL_1 was associated with all three non-AIDS events in a univariate analysis, but only associated with Hodgkin lymphoma and squamous cell carcinoma of the anus in a multivariate analysis.
Chriouze et al. 2015(11) APROCO-COPILOTE	**cVL_1;** Calculated by summing under the **viral load** curve **starting from 8 months post-baseline**	Cox proportional hazards model of **all-cause mortality** vs. dichotomized cVL_1. Adjusted for sex, age, ART	Baseline CD4 (at ART initiation): 278 (125-416) cVL_1 (dichotomized) was only statistically significantly associated with all-cause mortality when

France 1997-2010	(at baseline cohort included both pretreated and ART-naïve patients) and then taking the logarithm.	status at baseline, history of AIDS event, baseline and time-updated CD4, baseline and time-updated log viral load.	time-updated log viral load was excluded from the analysis.
Sempa et al. 2015 (Current article) IDI Cohort, Uganda 2004-2013	Calculated by summing under the **viral load** (cVL_1) or **log viral load** (cVL_2) curve (6 monthly measurements) and above the log (400cp/μl) detectability threshold **starting from ART initiation**; log (400) cutoff was chosen so that undetectable viral loads did not contribute to cumulative viremia. In a sensitivity analysis we used 1og (1000)	Cox proportional hazards model of **opportunistic infection**, **AIDS-related mortality**, or **all-cause mortality** vs. either cVL_1 or cVL_2. Adjusted for time-updated and baseline log viral load, time-updated and baseline CD4, baseline age and sex. Included interaction between laboratory measurements and time since measurement to include declining effect of measurement over time when outcomes observed more frequently than covariates.	Baseline CD4 (at ART initiation): 100 (38-168) Neither cVL measure was significantly associated with opportunistic infection risk, which was better predicted by time-updated viral load, hemoglobin levels and CD4 count. cVL_2, but neither cVL_2 nor time-updated log viral load, was significantly associated with mortality risk. Lower CD4 and lower hemoglobin were also significantly associated with increased mortality risk. Viral load measurements were only predictive of opportunistic infection or mortality risk for the 12 weeks post-measurement, while other variables were predictive of mortality (hemoglobin, CD4) or opportunistic infection (hemoglobin) risk for up to 24 weeks. cVL_2 remained a significant predictor

[a]Only covariates corresponding to viral load or CD4 measures are described in this table.
[b]We do not report hazard ratios because they are not directly comparable between studies that modeled viral loads and cumulative viremia calculated in different ways.
[d]All baseline CD4 given as median (IQR) except for [e].
[e]Breakdown of pre-ART CD4 nadir by category

Web Appendix

All-Cause Mortality Outcome

While we assessed all-cause mortality as an outcome, we censored one patient who died after a motor accident at their death, excluding this outcome *a priori* from the analysis since this was judged to be unrelated to exposure to HIV.

Cox Proportional Regression Equations

We used the following Cox proportional hazards model:

$$\log(\lambda_{i,t}) = H_0(t) + \boldsymbol{\beta}_{\text{lab}}\boldsymbol{X}_{\text{lab,i,t}} + \boldsymbol{\beta}_{\text{other}}\boldsymbol{X}_{other,i,t}$$

$$\boldsymbol{\beta}_{\text{lab}}\boldsymbol{X}_{\text{lab},i,t} = \sum_{\substack{\text{all } m \\ \text{laboratory} \\ \text{variables}}} \beta_m X_{m,i,t} + \beta_m^* X_{m,i,t} X_{\text{lag},i,t}$$

$$\boldsymbol{\beta}_{\text{other}}\boldsymbol{X}_{\text{other},i,t} = \sum_{\substack{\text{all } k \\ \text{other} \\ \text{variables}}} \beta_k X_{k,i,t}$$

for the hazard experienced by the *i*-th individual in the *t*-th time interval, allowing an interaction between laboratory measurements and the time since measurement $X_{lag,i,t}$ (governed by β^*). In this way, declining predictive utility of explanatory variables with time since measurement can be fit directly from the data. Laboratory variables included current and baseline viral load, current and baseline CD4, current hemoglobin and cVL_1 or cVL_2. Non-laboratory variables included current age and sex. We dealt with the detection threshold of viral load by using a categorical dummy variable and associated coefficients for undetectable viral load measurements, such that the regression terms for time-updated viral load in the model were;

$$(1 - X_{\text{undetVL},i,t})(\beta_{\text{VL}} X_{\text{VL},i,t} + \beta_{\text{VL}}^* X_{\text{VL},i,t} X_{\text{lag},i,t}) + \beta_{\text{undetVL}} X_{\text{undetVL},i,t}$$

where $X_{\text{undetVL,i,t}}$ is an indicator variable (1 if undetectable and 0 if detectable). Thus, for example, the adjusted relative hazard for a 1 $\log_{10}$ difference in viral load over the 0-12 weeks post-measurement is given by

$\text{ARH}_{\text{VL, 0-12 wk}} = \exp(\hat{\beta}_{VL} + \widehat{\beta^*}_{VL} \times 12)$

Wald Chi Squared Confidence Intervals

We used Wald confidence interval calculations to assess the significance of the relationship between each predictor and outcome, after controlling a number of other variables. As noted above, we have used two coefficients for laboratory measurement variables (and three for viral load due to the detectability dummy variable). Confidence intervals were constructed based on linear combinations of fitted coefficients using their variance covariance matrix and the appropriate variance transformations.

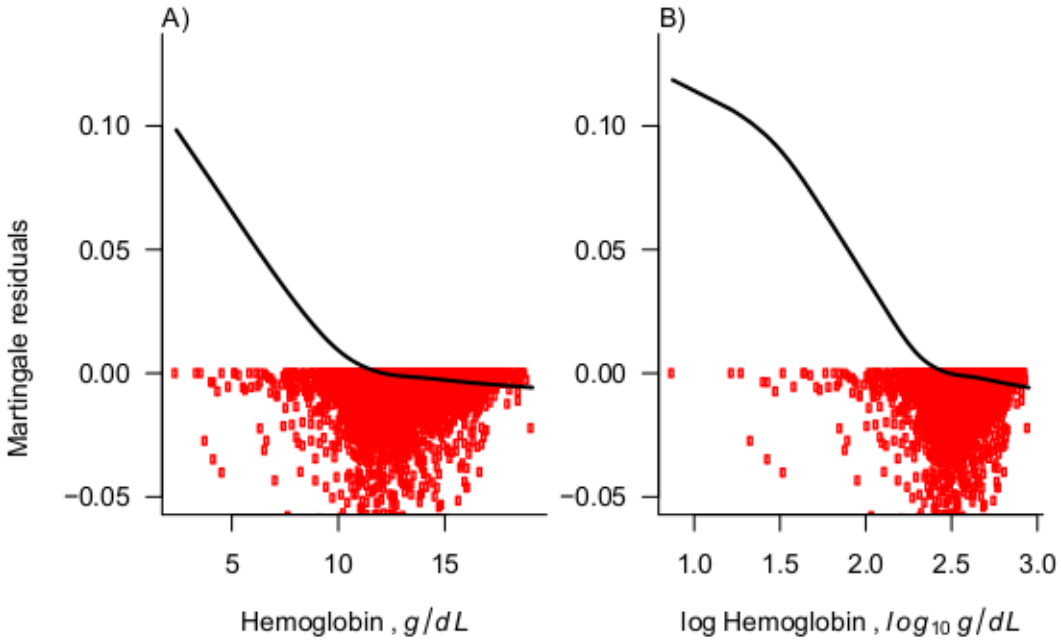
Cox Proportional Hazards Models and Rounded Observation Intervals

We simulated lognormal viral load data for the patient observation time points for the 489 patients from the IDI cohort, calculating time-updated cumulative viral load according to the methods in the main text. We then simulated incident OI times based on time-varying hazards and an assumed causal relationship between incident OI and viral load, but not cumulative log viral load. Using actual start and end times and visit intervals, we simulated the dataset (adding incident OIs, viral load and cVL) 1000-times. We analyzed the resulting simulated datasets (10,657 data points of 489 patients with an average of 1,162 OI events) with Cox proportional hazards model as in the main text and also with a Poisson regression model (cloglog link generalized linear mixed model; GLMM), using exact start and end points for each inter-visit observational interval. We found that results from the Cox proportional hazards model, but not from the Poisson GLMM to be biased. Specifically, Cox proportional hazards

models consistently found cVL to be significantly predictive of OI risk even when though was no underlying causal effect, while P-values for the association of cVL with the outcome were uniformly distributed from 0 to 1 for the Poisson GLMM (Web Figure 4). However, when we analyzed the simulations with a Cox proportional hazards model *in which inter-visit intervals were rounded* to their approximate 12 week values (as designated by the original study design), the analysis was unbiased, failing to spuriously attribute a significant association between cVL and OI risk more than the nominal $\alpha = 0.05$ false positive rate. We therefore used rounded inter-visit observation intervals when using the Cox proportional hazards model to analyze the IDI cohort data in the main text. Please follow this link: *https://Sempa@github.com/ICI3D/SempaetalAJE-00426-2015.git*, to view or run the R-file *"CoxPHbiasFromUnroundedObsTimes"* for simulation details.

Baseline regimen

Including baseline ART regimen (nevirapine or efavirenz based regimen) as covariates in survival models could cause confounding bias because at baseline patients were allocated to nevirapine or efavirenz based on aspects of clinical presentation that are already included via other covariates (1). To avert this situation, we used regression trees to generate propensity scores (2) using baseline variables: viral load, CD4 count, hemoglobin, ART regimen, age, and gender to adjust for bias in treatment allocation (nevirapine or efavirenz) at ART initiation (see Web Figure 5). After pruning—removing highly specific nodes—there was only one root, which implied that these variables were not informative with regard to treatment allocation. We therefore completed the analysis without using propensity scores or ART regimen.



Web Figure 1. Loess-smoothed Martingale Residuals for All-cause Mortality Outcomes versus Hemoglobin counts for HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013.

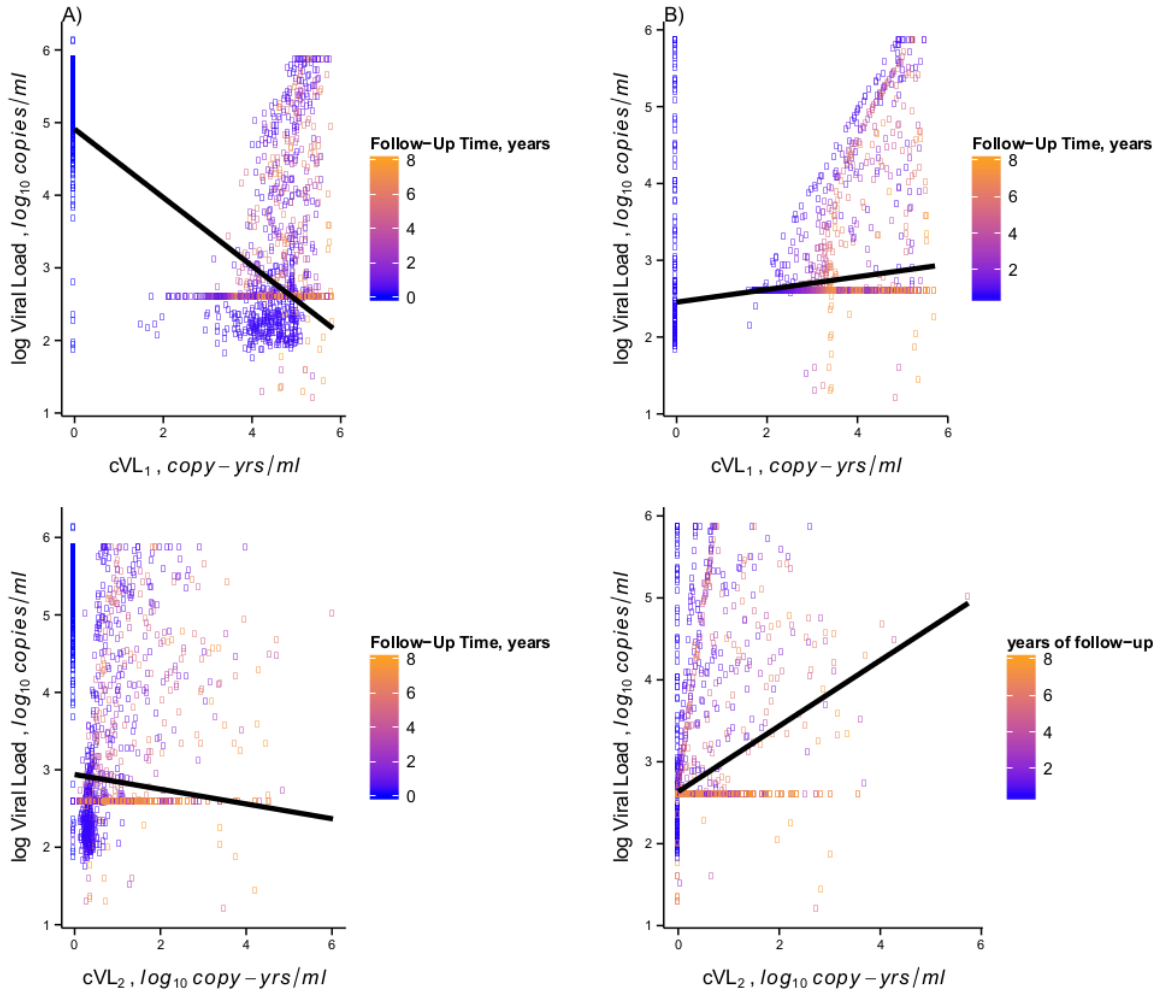
Multivariate Cox proportional hazards were fit with all variables as indicated in the main text except for time-updated hemoglobin. Martingale residuals were then plotted versus hemoglobin on a linear (Web Figure 1A) and logarithmic scale (Web Figure 1B), with a loess trend to visually inspect their functional relationship. The trend with hemoglobin on a log scale is better approximated by a linear relationship, justifying the inclusion of hemoglobin's inclusion in the model as log hemoglobin. Similar visual inspections were used to determine the specification of each covariate.

Web Table 2: Characteristics of the 489 for HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013 included in the analysis.

Variable	Median (IQR)
Baseline Age (years)	35.3 (30.2 – 41.8)
Gender: n (%)	
Female	341 (69.7)
Male	148 (30.3)
Baseline CD4 count (cells/μL)	100 (30-168)
nevirapine based regimen at baseline: n (%)	363 (74.2)
Baseline viral load : Log10 copies/ml	5.4 (5.1 – 5.8)
Follow-up time (years)	8.3 (2.3 – 8.8)

Web Table 3: Spearman Correlation Matrix between Viral Load and CD4 variables among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013. Correlations displayed include 11819 observations of 489 patients where variables indicate either time-varying measurements (log (VL), cVL_2, cVL_1, log (CD4)) or a single measurement for each patient (baseline log (VL), peak log (VL)).

	log (VL)	**cVL_1**	**cVL_2**	**baseline log (VL)**	**peak log (VL)**	**log (CD4)**
log (VL)	1	-0.69	-0.034	0.045	0.07	-0.49
cVL_1		1	0.47	0.32	0.34	0.47
cVL_2			1	0.15	0.22	0.094
baseline log (VL)				1	0.89	0.0049
peak log (VL)					1	-0.0094
log (CD4)						1



Web Figure 2: Correlation between Cumulative HIV-Viremia Metrics and log Viral Load for HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013.

Web Figure 2A— correlation between log VL, and cVL measures when baseline is at baseline visit. Web Figure 2B— correlation between log VL, and cVL measures when baseline is shifted to week 24. Each point shows a single laboratory measurement, with color indicating the time since ART initiation (i.e. years of follow-up) for that measurement. A linear model (black line) is displayed to illustrate the correlation shown in Web Table 3. The strong negative correlation between cVL_1 and log viral load is driven by cVL_1's rapid increases at the baseline

visit when viral load is high. Because accumulation on a linear scale means that cVL_1 only increases slightly for subsequent intermediate viral load measurements.

Web Table 4: Sensitivity analysis of Opportunistic Infection Model Results using different viral load detection thresholds among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013. Values give adjusted hazard ratios (95% confidence interval) for the hazard of acquiring an incident opportunistic infection from multivariate Cox proportional hazard models with cVL_2 calculated using either viral load detection thresholds of either 400 or 1000 copies/ml.

	Threshold for cVL_2 calculation			
	400 copies/ml		1000 copies/ml	
	AHR	95% CI	AHR	95% CI
per log10 increase in VL, *log_{10} copies/ml*				
predicting 0-12 weeks ahead	1.34	1.120, 1.610 [c]	1.35	1.130, 1.620 [b]
predicting 0-24 weeks ahead	1.21	0.969, 1.500	1.21	0.968, 1.500
per log10 increase in cumulative viremia				
predicting 0-12 weeks ahead	0.78	0.523, 1.150	0.72	0.442, 1.180
predicting 0-24 weeks ahead	1.00	0.679, 1.480	0.99	0.610, 1.600
per 2-fold increase in CD4 count, *cells/μL*				
predicting 0-12 weeks ahead	0.90	0.804, 0.998 [a]	0.90	0.803, 0.999 [a]
predicting 0-24 weeks ahead	0.91	0.755, 1.110	0.91	0.754, 1.110
per 10% increase in hemoglobin				
predicting 0-12 weeks ahead	0.91	0.859, 0.959 [c]	0.91	0.859, 0.959 [c]
predicting 0-24 weeks ahead	0.89	0.819, 0.971 [b]	0.89	0.818, 0.971 [b]
per 2-fold increase in baseline CD4 count, *cells/μL*	0.98	0.898, 1.080	0.98	0.898, 1.080
Baseline viral load, *log_{10} copies/ml*				
1st	1			1
2nd	0.96	0.692, 1.320	0.96	0.693, 1.320
3rd	1.20	0.869, 1.640	1.20	0.872, 1.650
4th	1.01	0.715, 1.420	1.01	0.718, 1.430
Gender				
Female	1			1

Male	0.78	0.602, 1.010	0.78	0.603, 1.010
per 10 year increase in baseline age	0.91	0.791, 1.040	0.91	0.791, 1.040

[a] Statistical significance: $P < 0.05$; [b] $P < 0.01$; [c] $P < 0.001$

Quartile: 1st—$<10^{5.07}$; **2nd**—$10^{5.08}$ - $10^{5.44}$; **3rd**—$10^{5.45}$ - $10^{5.77}$; **4th**—$10^{5.78}$ - $10^{6.15}$

ART—Antiretroviral therapy; **AHR**—Adjusted Hazard Ratio; **HIV**—Human Immune Virus; **cVL$_1$**—log cumulative Viral Load; **cVL$_2$**—cumulative log Viral Load; **VL**— Viral Load

Web Table 5: Sensitivity analysis of All-cause Mortality Model Results using different viral load detection thresholds among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013. Values give adjusted hazard ratios (95% confidence interval) for the hazard of dying of any cause from multivariate Cox proportional hazard models with cVL_2 calculated using either viral load detection thresholds of either 400 or 1000 copies/ml.

Variable	Threshold for cVL_2 calculation			
	400 copies/ml		1000 copies/ml	
	AHR	95% CI	AHR	95% CI
per log10 increase in VL, *log_{10} copies/ml*				
predicting 0-12 weeks ahead	1.13	0.722, 1.770	1.11	0.710, 1.740
predicting 0-24 weeks ahead	0.89	0.512, 1.550	0.85	0.491, 1.490
per log10 increase in cumulative viremia, *log_{10} copy-yrs/ml*				
predicting 0-12 weeks ahead	1.63	1.020, 2.600 [a]	1.86	1.060, 3.260 [a]
predicting 0-24 weeks ahead	0.50	0.168, 1.490	0.28	0.0623, 1.210
per 2-fold increase in CD4 count, *cells/μL*				
predicting 0-12 weeks ahead	0.57	0.454, 0.723 [c]	0.57	0.453, 0.720 [c]
predicting 0-24 weeks ahead	0.69	0.514, 0.922 [a]	0.69	0.517, 0.926 [a]
per 10% increase in hemoglobin, *g/dl*				
predicting 0-12 weeks ahead	0.77	0.702, 0.832 [c]	0.76	0.702, 0.831 [c]
predicting 0-24 weeks ahead	0.73	0.650, 0.817 [c]	0.73	0.654, 0.822 [c]
per 2-fold increase in baseline CD4 count, *cells/μL*	1.1	0.920, 1.280	1.08	0.920, 1.280
Baseline viral load, *log_{10} copies/ml*				
1st	1			1
2nd	1.51	0.682, 3.330	1.53	0.687, 3.390
3rd	1.28	0.527, 3.090	1.31	0.535, 3.220
4th	3.62	1.710, 7.640 [c]	3.71	1.740, 7.930 [c]
Gender				
Female	1			1
Male	1.07	0.556, 2.050	1.07	0.556, 2.050
Baseline age, *years*				

≤ 35	1			1
36 – 45	1.29	0.713, 2.340	1.31	0.719, 2.370
46 – 55	1.71	0.815, 3.600	1.74	0.828, 3.650
≥ 56	3.02	1.300, 6.970 [b]	3.09	1.340, 7.150 [b]

[a] Statistical significance: $P < 0.05$; [b] $P < 0.01$; [c] $P < 0.001$
Quartile: 1st—$<10^{5.07}$; **2nd**—$10^{5.08}$ - $10^{5.44}$; **3rd**—$10^{5.45}$ - $10^{5.77}$; **4th**—$10^{5.78}$ - $10^{6.15}$
ART—Antiretroviral therapy; **AHR**—Adjusted Hazard Ratio; **HIV**—Human Immune Virus; **cVL1**—log cumulative Viral Load; **cVL2**—cumulative log Viral Load; **VL**— Viral Load

Web Table 6: Sensitivity Analysis of Opportunistic Infection Model Results among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013. The sensitivity analysis involved recalculating cVL by moving baseline viral load from baseline visit to 24 week measurement. Values give adjusted hazard ratios (95% confidence interval) for the hazard of acquiring an incident opportunistic infection from multivariate Cox proportional hazard models with cumulative viremia calculated as one of either cVL_1 or cVL_2.

Variable	Model with cVL_1		Model with cVL_2	
	AHR	95% CI	AHR	95% CI
per log10 increase in VL, *log_{10} copies/ml*				
predicting 0-12 weeks ahead	1.30	1.030, 1.640 [a]	1.30	1.030, 1.640 [a]
predicting 0-24 weeks ahead	1.57	1.130, 2.180 [b]	1.53	1.090, 2.150 [a]
per log10 increase in cumulative viremia, *log_{10} copy-yrs/ml*				
predicting 0-12 weeks ahead	0.91	0.689, 1.190	0.76	0.416, 1.400
predicting 0-24 weeks ahead	0.89	0.659, 1.190	0.86	0.451, 1.630
per 2-fold increase in CD4 count, *cells/µL*				
predicting 0-12 weeks ahead	0.88	0.635, 1.230	0.89	0.639, 1.230
predicting 0-24 weeks ahead	0.98	0.645, 1.500	0.97	0.646, 1.460
per 10% increase in hemoglobin, *g/dl*				
predicting 0-12 weeks ahead	0.91	0.840, 0.987 [a]	0.91	0.838, 0.981 [a]
predicting 0-24 weeks ahead	0.85	0.749, 0.962 [a]	0.85	0.751, 0.966 [a]
per 2-fold increase in baseline CD4 count, *cells/µL*	1.01	0.716, 1.420	1.01	0.715, 1.420
Baseline viral load, *log_{10} copies/ml*				
1st	1			1
2nd	0.87	0.551, 1.380	0.87	0.550, 1.370
3rd	1.10	0.705, 1.720	1.10	0.705, 1.720
4th	0.86	0.539, 1.360	0.85	0.537, 1.350
Gender				
Female	1			1
Male	0.88	0.588, 1.310	0.88	0.590, 1.310

per 10 year increase in baseline age	1.10	0.917, 1.330	1.10	0.916, 1.330

[a] Statistical significance: $P < 0.05$; [b] $P < 0.01$; [c] $P < 0.001$
Quartile: 1st—$<10^{5.07}$; **2nd**—$10^{5.08}$ - $10^{5.44}$; **3rd**—$10^{5.45}$ - $10^{5.77}$; **4th**—$10^{5.78}$ - $10^{6.15}$
ART—Antiretroviral therapy; **AHR**—Adjusted Hazard Ratio; **HIV**—Human Immune Virus; **cVL1**—log cumulative Viral Load; **cVL2**—cumulative log Viral Load; **VL**— Viral Load

Web Table 7: Sensitivity Analysis of All-cause Mortality Model Results among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013. The sensitivity analysis involved recalculating cVL by moving baseline viral load from baseline visit to 24 week measurement. Values give adjusted hazard ratios (95% confidence interval) for the hazard of dying of any cause from multivariate Cox proportional hazard models with cumulative viremia calculated either as cVL_1 or cVL_2.

Variable	Model with cVL_1		Model with cVL_2	
	AHR	95% CI	AHR	95% CI
per log10 increase in VL, *log_{10} copies/ml*				
predicting 0-12 weeks ahead	0.93	0.529, 1.630	1.20	0.671, 2.160
predicting 0-24 weeks ahead	1.17	0.684, 2.010	1.32	0.768, 2.270
per log10 increase in cumulative viremia, *log_{10} copy-yrs/ml*				
predicting 0-12 weeks ahead	1.81	1.270, 2.580 [b]	1.81	1.010, 3.230 [a]
predicting 0-24 weeks ahead	0. 90	0.601, 1.350	0.58	0.219, 1.520
per 2-fold increase in CD4 count, *cells/µL*				
predicting 0-12 weeks ahead	0.54	0.399, 0.733 [c]	0.55	0.402, 0.738 [c]
predicting 0-24 weeks ahead	0.67	0.410, 1.110	0.61	0.378, 0.980 [a]
per 10% increase in hemoglobin, *g/dl*				
predicting 0-12 weeks ahead	0.78	0.718, 0.856 [c]	0.79	0.718, 0.862 [c]
predicting 0-24 weeks ahead	0.74	0.634, 0.868 [c]	0.73	0.633, 0.848 [c]
per 2-fold increase in baseline CD4, *cells/µL* **count**	1.00	0.710, 1.400	0.93	0.660, 1.310
Baseline viral load, *log_{10} copies/ml*				
1st	1			1
2nd	1.50	0.630, 3.580	1.59	0.648, 3.900
3rd	1.09	0.416, 2.860	1.14	0.428, 3.020
4th	2.46	1.100, 5.490 [a]	2.73	1.200, 6.230 [a]
Gender				
Female	1			1
Male	1.06	0.495, 2.260	1.08	0.488, 2.390

Baseline age, *years*				
≤ 35	1			1
36 – 45	1.61	0.737, 3.510	1.68	0.752, 3.760
46 – 55	1.94	0.765, 4.920	1.90	0.742, 4.850
≥ 56	3.98	1.430, 11.100 [b]	3.88	1.390, 10.800 [b]

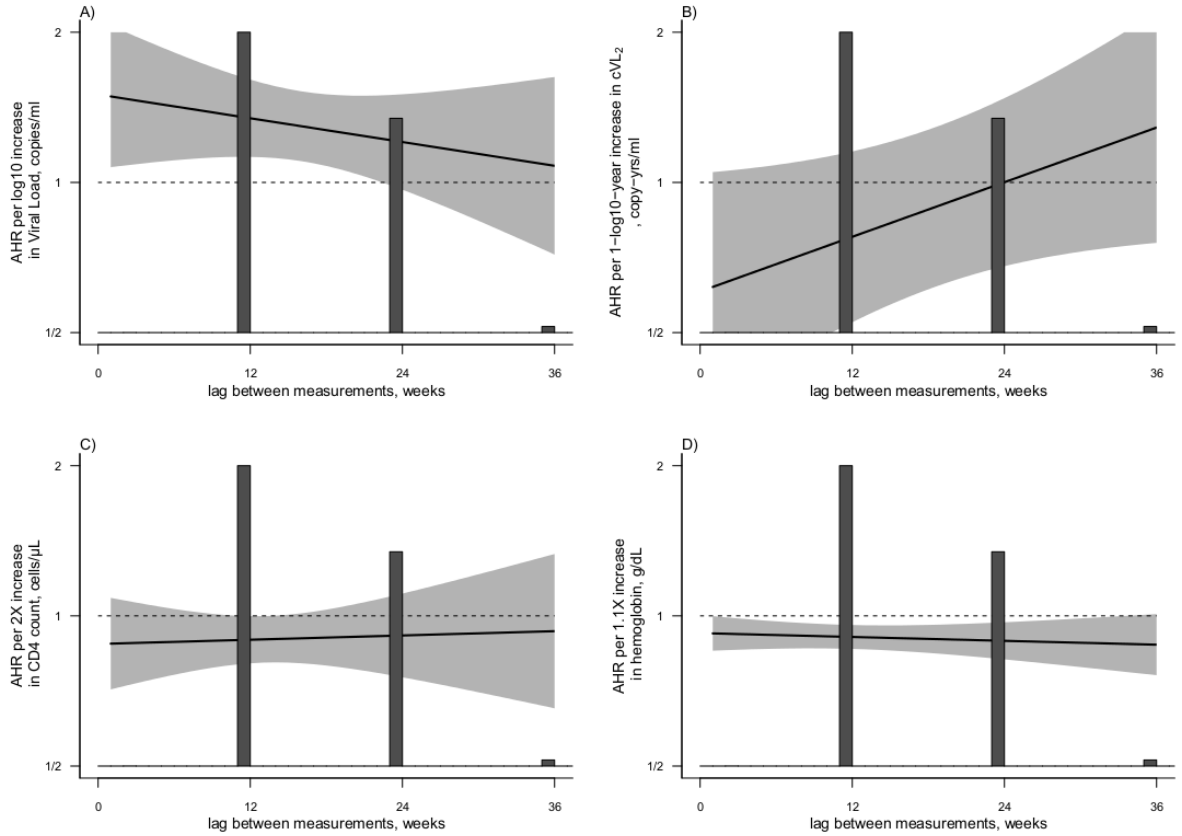
[a] Statistical significance: $P < 0.05$; [b] $P < 0.01$; [c] $P < 0.001$
Quartile: 1st—$<10^{5.07}$; **2nd**—$10^{5.08}$ - $10^{5.44}$; **3rd**—$10^{5.45}$ - $10^{5.77}$; **4th**—$10^{5.78}$ - $10^{6.15}$
ART—Antiretroviral therapy; **AHR**—Adjusted Hazard Ratio; **HIV**—Human Immune Virus; **cVL$_1$**—log cumulative Viral Load; **cVL$_2$**—cumulative log Viral Load; **VL**— Viral Load

Web Table 8: HIV specific Mortality Model Results among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013. Values give adjusted hazard ratios (95% confidence interval) for the hazard of HIV-related causes from multivariate Cox proportional hazard models with cumulative viremia calculated either as cVL_1 or cVL_2.

Variable	Model with cVL_1		Model with cVL_2	
	AHR	95% CI	AHR	95% CI
per log10 increase in VL, *log_{10} copies/ml*				
predicting 0-12 weeks ahead	1.05	0.580, 1.900	0.94	0.464, 1.920
predicting 0-24 weeks ahead	1.58	0.789, 3.150	0.93	0.446, 1.960
per log10 increase in cumulative viremia, *log_{10} copy-yrs/ml*				
predicting 0-12 weeks ahead	1.18	0.493, 2.820	1.34	0.601, 2.980
predicting 0-24 weeks ahead	1.74	1.000, 3.030 [a]	1.20	0.415, 3.440
per 2-fold increase in CD4 count, *cells/µL*				
predicting 0-12 weeks ahead	0.60	0.429, 0.831 [b]	0.58	0.416, 0.798 [c]
predicting 0-24 weeks ahead	0.65	0.435, 0.969 [a]	0.72	0.489, 1.060
per 10% increase in hemoglobin, *g/dl*				
predicting 0-12 weeks ahead	0.71	0.637, 0.792 [c]	0.71	0.632, 0.786 [c]
predicting 0-24 weeks ahead	0.59	0.459, 0.760 [c]	0.63	0.509, 0.777 [c]
per 2-fold increase in baseline CD4 count , *cells/µL*	1.00	0.822, 1.220	1.01	0.822, 1.230
Baseline viral load, *log_{10} copies/ml*				
1st	1			1
2nd	2.19	0.694, 6.890	2.24	0.754, 6.630
3rd	1.95	0.528, 7.200	2.13	0.646, 7.040
4th	5.93	1.840, 19.200 [b]	6.99	2.580, 18.900 [c]
Gender				
Female	1			1
Male	0.87	0.335, 2.250	0.85	0.321, 2.250
Baseline age, *in years*				
≤ 35	1			1
36 – 45	1.65	0.782, 3.480	1.58	0.719, 3.460

46 – 55	1.32	0.428, 4.060	1.33	0.432, 4.110
≥ 56	2.29	0.563, 9.340	2.32	0.599, 8.980

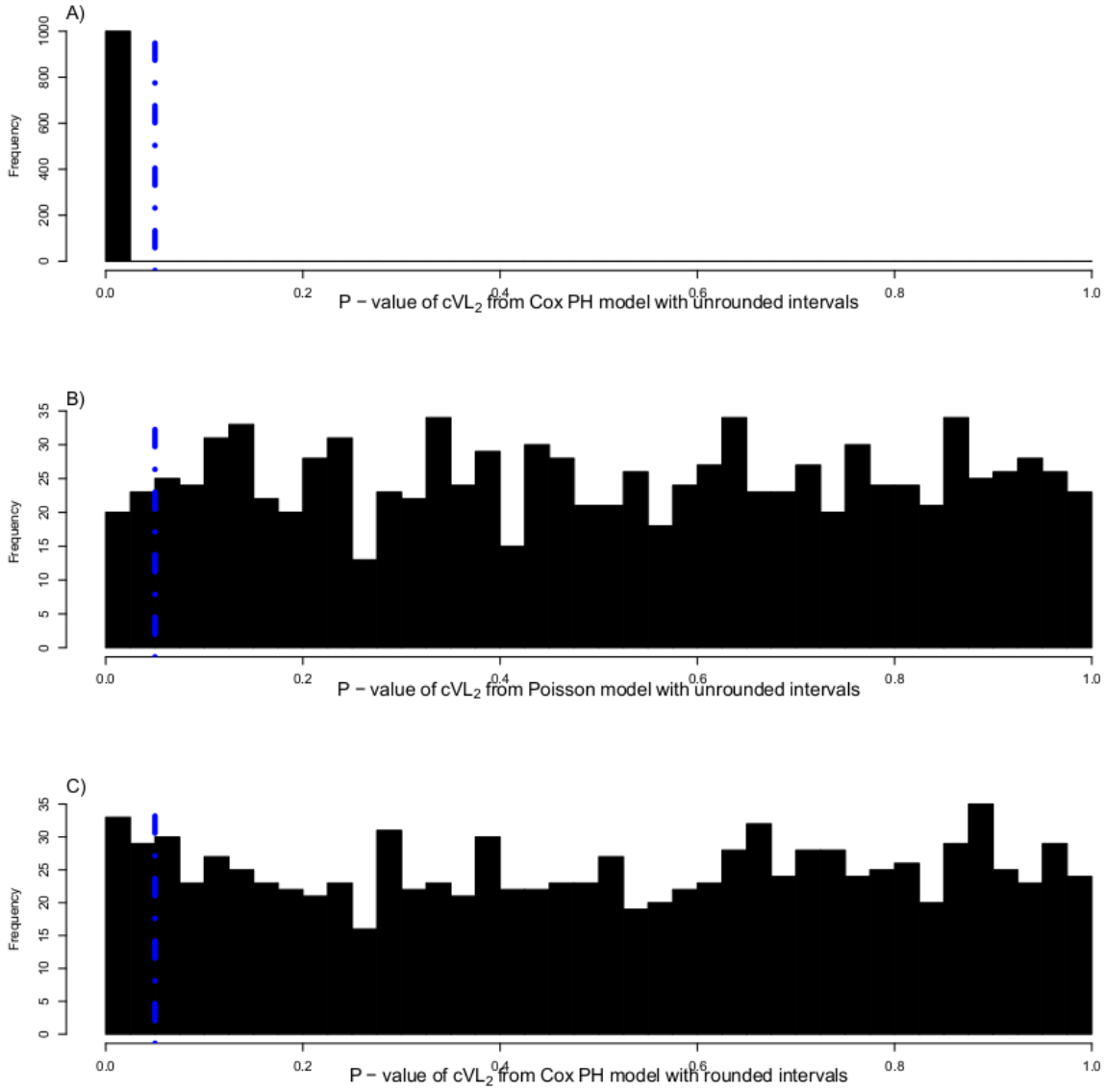
[a] Statistical significance: $P < 0.05$; [b] $P < 0.01$; [c] $P < 0.001$
Quartile: 1st—$<10^{5.07}$; **2nd**—$10^{5.08}$ - $10^{5.44}$; **3rd**—$10^{5.45}$ - $10^{5.77}$; **4th**—$10^{5.78}$ - $10^{6.15}$
ART—Antiretroviral therapy; **AHR**—Adjusted Hazard Ratio; **HIV**—Human Immune Virus; **cVL$_1$**—log cumulative Viral Load; **cVL$_2$**—cumulative log Viral Load; **VL**— Viral Load



Web Figure 3: Declining Prognostic Value with Increasing Time since Measurement among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013.

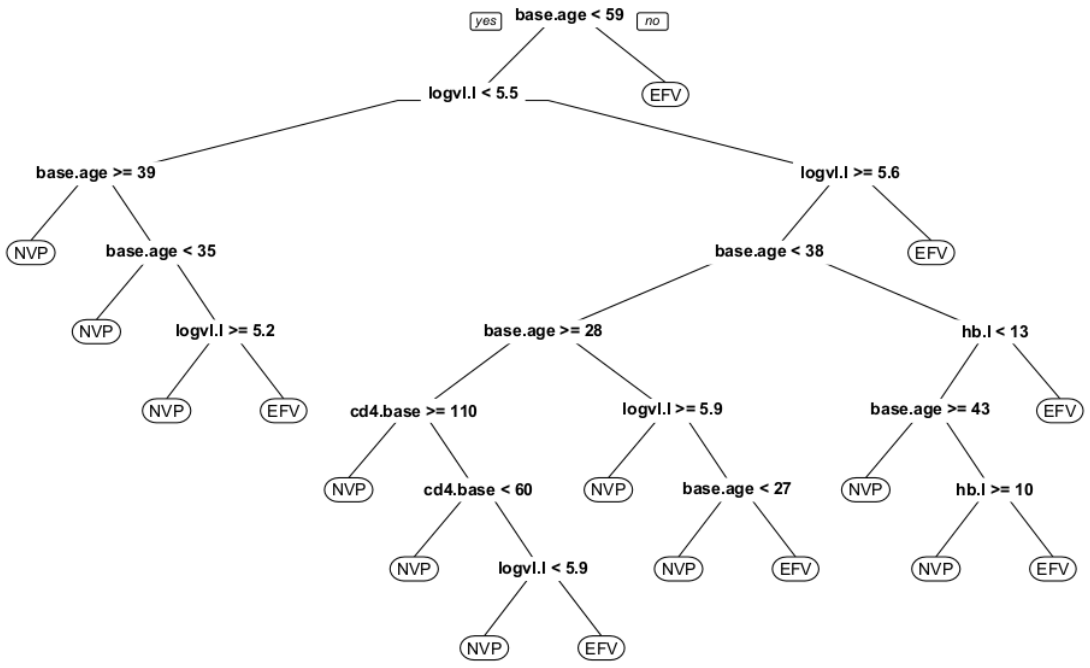
Lines and shaded regions show the adjusted hazard ratio (AHR) and 95% confidence intervals, respectively, for the hazard of acquiring an incident OI per log10 increase in viral load (Web Figure 3A), per log10 increase in cVL (Web Figure 3B), per 50 increase in CD4 count (Web Figure 3C) and per 1g/dl increase in hemoglobin (Web Figure 3D). The AHR is modeled as a function of the time since the laboratory measurement was made, facilitating the estimation of how the predictive utility of each measurement declines over time. Inset histograms show the frequency distribution of time lags between a clinic visit and the time of the last laboratory measurement. Because most patients visited quarterly but only had

laboratory assays performed every other visit, most time lags were at 12 and 24 weeks and visits are rounded to 12-week intervals in the analysis.



Web Figure 4: Distribution of P-values for the effect of cVL on OI risk amongst 1000 simulations assuming no actual effect.

The graphs are in the order Cox proportional hazards model with unrounded inter-visit intervals, Poisson regression with unrounded inter-visit intervals and Cox proportional hazards model with rounded inter-visit intervals.



Web Figure 5: Probability of receiving nevirapine or efavirenz among HIV Patients on ART in the IDI cohort, Kampala, Uganda, 2004-2013.

Where NVP - nevirapine; EFV- efavirenz; base.age – Age at baseline; cd4.base – Baseline CD4 count, logv.l – Baseline HIV log viral load; hb.l – Baseline hemoglobin

References

1. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu. Rev. Public Health.* 2000;21:121–145.

2. Westreich D, Lessler J, Funk MJ. Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *J. Clin. Epidemiol.* [electronic article]. 2010;63(8):826–33. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2907172&tool=pmcentrez&rendertype=abstract). (Accessed July 15, 2014)

3. Cole SR, Napravnik S, Mugavero MJ, et al. Copy-years viremia as a measure of cumulative human immunodeficiency virus viral burden. *Am. J. Epidemiol.* [electronic article]. 2010;171(2):198–205. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2878100&tool=pmcentrez&rendertype=abstract). (Accessed February 9, 2014)

4. Zoufaly A, Stellbrink H-J, Heiden MA der, et al. Cumulative HIV viremia during highly active antiretroviral therapy is a strong predictor of AIDS-related lymphoma. *J. Infect. Dis.* [electronic article]. 2009;200(1):79–87. (http://www.ncbi.nlm.nih.gov/pubmed/19476437). (Accessed September 3, 2014)

5. Marconi VC, Grandits G, Okulicz JF, et al. Cumulative viral load and virologic decay patterns after antiretroviral therapy in HIV-infected subjects influence CD4 recovery and AIDS. *PLoS One* [electronic article]. 2011;6(5):e17956. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3098832&tool=pmcentrez&rendertype=abstract). (Accessed January 21, 2014)

6. Mugavero MJ, Napravnik S, Cole SR, et al. Viremia copy-years predicts mortality among treatment-naive HIV-infected patients initiating antiretroviral therapy. *Clin. Infect. Dis.* [electronic article]. 2011;53(9):927–935. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3189165&tool=pmcentrez&rendertype=abstract). (Accessed March 3, 2014)

7. Mugavero MJ, Amico KR, Westfall AO, et al. Early retention in HIV care and viral load suppression: implications for a test and treat approach to HIV prevention. *J. Acquir. Immune Defic. Syndr.* [electronic article]. 2012;59(1):86–93. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3237801&tool=pmcentrez&rendertype=abstract)

8. Saracino A, Monno L, Scudeller L, et al. X4 viruses are frequently archived in patients with long-term HIV infection but do not seem to influence the “inflamm-aging” process. *BMC Infect. Dis.* [electronic article]. 2013;13(1):220. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3661370&tool=pmcentrez&rendertype=abstract)

9. Lima VD, Sierra-madero J, Wu Z, et al. Comparing the efficacy of efavirenz and boosted lopinavir using viremia copy-years. *J. Int. AIDS Soc.* 2014;17(18617):1–4.

10. Kowalkowski M, Day R, Du X, et al. Cumulative Hiv Viremia and Non-Aids-Defining Malignancies Among a Sample of Hiv-Infected Male Veterans. *J. Acquir. Immune Defic. Syndr.* [electronic article]. 2014;67(2):204–211. (http://www.ncbi.nlm.nih.gov/pubmed/25078536)

11. Chirouze C, Journot V, Moing V Le. Viremia Copy-Years as a Predictive Marker of All-Cause Mortality in HIV-1 – Infected Patients Initiating a Protease Inhibitor – Containing Antiretroviral Treatment. *J. Acquir. Immune Defic. Syndr.* [electronic article]. 2015;68(2):204–208. (http://www.ncbi.nlm.nih.gov/pubmed/25590273)