APPENDIX

APPENDIX METHODS

Comparison of models and selection of final model using Akaike's information criterion (AIC)
As we indicated in the Methods section, "A smaller AIC indicates a better model; a difference in AIC of >10 between models is considered meaningful." The comparisons of models were made in relation to the model with the smallest AIC. In comparisons among models containing single measures, there was always a clear "winner," with one exception described below. When two models, with at least one of the models including more than one variable, had similar AICs, we chose the more parsimonious model. For example, for the final model for overall non-Hodgkin lymphoma (NHL) risk, we favored the model containing recent CD4 and late HIV RNA average (AIC=6488) over the model containing recent CD4, late HIV RNA average, and HIV RNA lagged by 540 days (AIC=6486), even though the latter model had a slightly lower AIC.

In comparisons of models containing single measures, there was only one instance in which the model with the smallest AIC and another model had similar AICs (a difference <8). In the Burkitt lymphoma (BL)-specific analysis (details not shown in the manuscript), the AIC for HIV RNA lagged by 900 days (613) was similar to the AIC for the late proportion of time HIV RNA >500 copies per mL (609), the HIV RNA measure with the lowest AIC. Adding HIV RNA lagged by 900 days to the model with late proportion of time HIV RNA >500 copies per mL did not improve the model (AIC remained at 609) and only the association for late proportion of time HIV RNA >500 copies per mL remained statistically significant. Furthermore, the lowest AIC for a CD4 measure was 656 and none of the best CD4 measures improved the fit when added to the model with late proportion of time HIV RNA >500 copies per mL. We therefore selected late proportion of time HIV RNA >500 copies per mL as the sole key predictor for BL risk.

Testing for subtype heterogeneity
To test whether the effects of the key CD4 and HIV RNA predictors for overall NHL risk varied (i.e., were heterogeneous) across the five NHL subtypes (i.e., event-types), we used a competing risk approach that required augmenting the dataset five times (i.e., one for each event-type) and using an interaction version of a stratified Cox model where the strata were the event-types.[1] We fit an event-stratified Cox model which, as the simple Cox model, was adjusted for cohort by stratification and for demographics (i.e., sex, race/ethnicity, and baseline age and calendar period) by multivariable adjustment. We estimated associations between the key predictors for overall NHL risk and risk for each NHL subtype, and then tested for heterogeneity of those associations (p-heterogeneity) by including product (i.e., interaction) terms between those key predictors and the event-types.

Imputation of unknown values and sensitivity analyses
For the primary analyses, we imputed race/ethnicity for persons with unknown race/ethnicity (5.5%) using cohort-specific probability weights based on sex and baseline age and calendar year. We also imputed HIV risk group (injection drug use, men who have sex with men, heterosexual, other) and smoking status (ever, never) for persons with unknown values, except for cohorts with a high proportion of unknowns, or, for smoking, with all the knowns being smokers. However, because of the large number of unknowns for HIV risk group and smoking status, we did not adjust for these two variables in the primary analysis but did so in sensitivity analyses, either using multiple imputation or treating unimputed unknowns as a separate category. To perform the multiple imputation, we first imputed (five times to generate five complete datasets) HIV risk group via the same procedure as for race/ethnicity, and smoking status via logistic regression for monotone missing data with the following set of covariates: cohort, sex, race/ethnicity, ever on antiretroviral therapy, follow-up time in the participant cohort, vital status, and baseline age, calendar period, HIV risk group, CD4, and HIV RNA. For HIV risk group and smoking status separately, we then ran the final model adjusting for the imputed variable using each imputed dataset. Next, using SAS PROC MIANALYZE, we combined the results from each complete dataset to produce pooled point estimates (i.e., average of the point estimates) and their respective standard errors adjusted to account for the imputation.[2] In a separate sensitivity analysis, we adjusted our final model for proportion of time on antiretroviral therapy during a moving window from 1260 days to 180 days in the past (determined by the key predictors included in the final model). In these sensitivity analyses, we assessed whether hazard ratios (HRs) meaningfully changed toward the null (>10%) after adjustment.

APPENDIX RESULTS

Selection of model for overall NHL risk

Comparison of separate models for recent, past, cumulative, and nadir/peak measures among persons with follow-up >2340 days (N=45,108; NHL cases=217)

For both CD4 and HIV RNA, lagged measures ≤1260 days in the past were better predictors (i.e., had lower AICs) than lagged measures >1260 days in the past (Appendix Table 3). For CD4, with increasing lag, the strength of the association with overall NHL risk decreased gradually, and the magnitude of AICs increased gradually. For HIV RNA, both the highest HRs and the lowest AIC occurred at the 540-day lag, followed by a graded HR decrease and AIC increase. Cumulative and nadir/peak measures during the "late" or more recent past (i.e., 1260 to 180 days [~3·5 years to ~6 months] in the past) were also better predictors than cumulative and nadir/peak measures during the "early" or more distant past (i.e., 2340 to 1260 days [~6·5 to ~3·5 years] in the past) or during the "early" and "late" past combined (i.e., 2340 to 180 days [~6·5 years to ~6 months] in the past; "overall"; Appendix Table 4). Based on these findings, we decided to focus on comparisons of models for measures occurring ≤1260 days in the past among the larger set of persons with follow-up >1260 days (N=68,585; NHL cases=403).

Comparison of separate models for recent, past, cumulative, and nadir/peak measures among persons with follow-up >1260 days (N=68,585; NHL cases=403)

Among recent and lagged CD4 measures, recent CD4 (i.e., lagged by 180 days) showed the highest HRs and the lowest AIC (6625), indicating it was the best CD4 measure from among the lagged CD4 measures (Appendix Table 3). Among recent and lagged HIV RNA measures, HIV RNA lagged by 540 days showed the highest HRs and the lowest AIC (6582), indicating it was the best HIV RNA measure from among the lagged HIV RNA measures (Appendix Table 3). HIV RNA lagged by 540 days remained superior to the recent HIV RNA (i.e., lagged by 180 days) when comparing AICs in models among persons with follow-up >540 days (a larger set of persons; data not shown).

Among cumulative and nadir/peak CD4 measures, the AICs for late CD4 average and late proportion of time CD4 <200 cells per μL were similar (6663 vs. 6664) and lower than the AIC for late CD4 nadir (6673; Appendix Table 4). Among cumulative and nadir/peak HIV RNA measures, late HIV RNA average had a lower AIC than late proportion of time HIV RNA >500 copies per mL (6522 vs. 6619) or late HIV RNA peak (6596); and, of all HIV RNA cumulative and nadir/peak measures, showed the strongest association with overall NHL risk (Appendix Table 4).

Among all measures (recent, past, cumulative or nadir/peak), comparison of AICs showed recent CD4 (i.e., lagged by 180 days) to be the best CD4 predictor of overall NHL risk, and late HIV RNA average (i.e., from ~3·5 years to ~6 months in the past) to be the best HIV RNA predictor of overall NHL risk (Appendix Tables 3 and 4). We chose recent CD4 and late HIV RNA average along with the following measures for further consideration: late CD4 average, late proportion of time CD4 <200 cells per μL, late CD4 nadir, HIV RNA lagged by 540 days, late proportion of time HIV RNA >500 copies per mL, and late HIV RNA peak.

Selection of best CD4 and HIV RNA predictor(s), respectively

Adding various CD4 measures, one at a time, to a model with recent CD4 (the best CD4 predictor from above) did not meaningfully improve the model. Furthermore, when included in a model with recent CD4, each cumulative and nadir CD4 measure lost its significance, while recent CD4 remained highly significant (Appendix Table 5). Therefore, we only chose recent CD4 as a candidate for our final model.

When we added various HIV RNA measures, one at a time, to a model with late HIV RNA average (the best HIV RNA predictor from above), the only measure that improved the model was HIV RNA lagged by 540 days (AIC, 6513 vs. 6522; Appendix Table 5). Although this improvement was only borderline meaningful according to the criterion that an AIC difference >10 is meaningful, we conservatively selected HIV RNA lagged by 540 days, along with late HIV RNA average, as candidates for our final model.

APPENDIX REFERENCES

1 Kleinbaum DG, Klein M. Survival analysis. A self-learning text, third edition. New York, NY: Springer; 2012.
2 Rubin DB. Multiple imputation for nonresponse in surveys. New York, NY: John Wiley & Sons; 1987.

APPENDIX TABLES

Appendix Table 1. Non-Hodgkin lymphoma (NHL) classification scheme

NHL subtypes	Comprising NHL histologies
Overall NHL	Any NHL histology.
Central nervous system NHL (CNS-NHL)	Any NHL histology occurring in the CNS.
Systemic NHL subtypes	Any NHL histology not occurring in the CNS.
Diffuse large B-cell lymphoma (DLBCL)	Diffuse large B-cell lymphoma, primary effusion lymphoma, plasmablastic lymphoma, and large B-cell lymphoma arising in HHV8-associated multicentric Castleman disease.
Burkitt lymphoma (BL)	Burkitt lymphoma/leukemia.
Other specified NHL	Other specified B-cell and T-cell lymphomas/lymphocytic leukemias (e.g., follicular lymphoma, chronic lymphocytic leukemia/small lymphocytic lymphoma, lymphoplasmacytic lymphoma, Walderston macroglobulinemia, mantle cell lymphoma, marginal zone lymphoma, peripheral T-cell lymphoma, anaplastic large cell lymphoma, mycosis fungoides, Sezary syndrome, hepatosplenic T-cell lymphoma, NK/T cell lymphoma, and other specified lymphocytic leukemia/lymphoma).
NHL not otherwise specified (NHL-NOS)	Unspecified malignant lymphoma/lymphoid leukemia.

Appendix Table 2. Baseline characteristics of subjects diagnosed with NHL, NA-ACCORD, 1996-2014.

Characteristic	Central nervous system NHL (CNS-NHL) (N=67) N (%)	Diffuse large B-cell lymphoma (DLBCL) (N=358) N (%)	Burkitt lymphoma (BL) (N=83) N (%)	Other specified NHL (N=94) N (%)	NHL not otherwise specified (NHL-NOS) (N=110) N (%)
Sex					
Male	59 (88·1)	324 (90·5)	76 (91·6)	84 (89·4)	97 (88·2)
Female	8 (11·9)	34 (9·5)	7 (8·4)	10 (10·6)	13 (11·8)
Race/ethnicity					
Black	28 (41·8)	118 (33·0)	20 (24·1)	27 (28·7)	43 (39·1)
White	24 (35·8)	197 (55·0)	51 (61·4)	58 (61·7)	55 (50·0)
Hispanic	7 (10·4)	20 (5·6)	7 (8·4)	3 (3·2)	4 (3·6)
Other	4 (6·0)	11 (3·1)	2 (2·4)	4 (4·3)	5 (4·5)
Unknown imputed	4 (6·0)	12 (3·4)	3 (3·6)	2 (2·1)	3 (2·7)
Age, years					
18–29	12 (17·9)	26 (7·3)	8 (9·6)	3 (3·2)	5 (4·5)
30–39	25 (37·3)	107 (29·9)	24 (28·9)	20 (21·3)	23 (20·9)
40–49	21 (31·3)	136 (38·0)	36 (43·4)	37 (39·4)	53 (48·2)
≥50	9 (13·4)	89 (24·9)	15 (18·1)	34 (36·2)	29 (26·4)
Calendar period					
1996–1999	44 (65·7)	198 (55·3)	28 (33·7)	58 (61·7)	72 (65·5)
2000–2003	16 (23·9)	87 (24·3)	22 (26·5)	25 (26·6)	24 (21·8)
2004–2007	6 (9·0)	45 (12·6)	23 (27·7)	7 (7·4)	9 (8·2)
2008–2011	1 (1·5)	27 (7·5)	10 (12·0)	2 (2·1)	4 (3·6)
2012–2014	0 (0·0)	1 (0·3)	0 (0·0)	2 (2·1)	1 (0·9)
Combination ART naive					
No	18 (26·9)	96 (26·8)	18 (21·7)	23 (24·5)	25 (22·7)
Yes	49 (73·1)	262 (73·2)	65 (78·3)	71 (75·5)	85 (77·3)
CD4 count, cells per μL					
<50	31 (46·3)	42 (11·7)	6 (7·2)	10 (10·6)	17 (15·5)
50–<100	9 (13·4)	42 (11·7)	4 (4·8)	11 (11·7)	14 (12·7)
100–<200	11 (16·4)	82 (22·9)	16 (19·3)	14 (14·9)	20 (18·2)
200–<350	7 (10·4)	88 (24·6)	9 (10·8)	17 (18·1)	31 (28·2)
350–<500	4 (6·0)	48 (13·4)	25 (30·1)	24 (25·5)	14 (12·7)
≥500	5 (7·5)	56 (15·6)	23 (27·7)	18 (19·1)	14 (12·7)
HIV RNA level, copies per mL					
≤500	2 (3·0)	55 (15·4)	9 (10·8)	27 (28·7)	24 (21·8)
>500–<10,000	16 (23·9)	60 (16·8)	19 (22·9)	16 (17·0)	17 (15·5)
10,000–<100,000	20 (29·9)	142 (39·7)	32 (38·6)	25 (26·6)	40 (36·4)
≥100,000	29 (43·3)	101 (28·2)	23 (27·7)	26 (27·7)	29 (26·4)
HIV risk group					
Injection drug use	10 (14·9)	45 (12·6)	10 (12·0)	12 (12·8)	10 (9·1)
Men who have sex with men	17 (25·4)	120 (33·5)	33 (39·8)	21 (22·3)	15 (13·6)
Heterosexual	10 (14·9)	33 (9·2)	12 (14·5)	13 (13·8)	6 (5·5)
Other	2 (3·0)	9 (2·5)	0 (0·0)	0 (0·0)	5 (4·5)
Unknown imputed[a]	25 (37·3)	132 (36·9)	23 (27·7)	46 (48·9)	68 (61·8)
Unknown not imputed[a]	3 (4·5)	19 (5·3)	5 (6·0)	2 (2·1)	6 (5·5)
Smoking status					
Ever	32 (47·8)	223 (62·3)	52 (62·7)	60 (63·8)	68 (61·8)
Never	8 (11·9)	69 (19·3)	14 (16·9)	14 (14·9)	16 (14·5)

Unknown imputed[a]	12 (17·9)	28 (7·8)	9 (10·8)	7 (7·4)	4 (3·6)
Unknown not imputed[a]	15 (22·4)	38 (10·6)	8 (9·6)	13 (13·8)	22 (20·0)

ART=antiretroviral therapy. NHL=non-Hodgkin lymphoma.

[a] We imputed HIV risk group and smoking status for persons with unknown values, except for cohorts with a high proportion of unknowns, or, for smoking, with all the knowns being smokers.

Appendix Table 3. Recent and past CD4 count and HIV RNA level measures, and overall NHL risk, NA-ACCORD, 1996-2014.

Measures	Recent measures[a]		Past measures[a]					
	NHL cases	HR (95% CI)	NHL cases	HR (95% CI)	NHL cases	HR (95% CI)	NHL cases	HR (95% CI)
CD4 count, cells per µL		**180 d lag**		**540 d lag**		**900 d lag**		**1260 d lag**
<50	114	11·1 (8·6–14·4)	74	8·4 (6·3–11·4)	47	5·7 (4·1–8·1)	33	5·1 (3·4–7·6)
50–<100	66	5·8 (4·3–7·9)	60	5·8 (4·2–7·9)	11	4·3 (3·0–6·1)	34	4·1 (2·8–6·1)
100–<200	117	3·5 (2·7–4·5)	112	3·5 (2·7–4·6)	95	3·0 (2·3–4·0)	75	2·9 (2·1–4·0)
200–<350	150	2·1 (1·7–2·7)	124	1·8 (1·4–2·4)	107	1·6 (1·2–2·1)	94	1·7 (1·3–2·3)
350–<500	128	1·7 (1·3–2·1)	103	1·4 (1·1–1·9)	90	1·3 (1·0–1·7)	74	1·3 (1·0–1·8)
≥500	137	1·0 (ref)	124	1·0 (ref)	117	1·0 (ref)	93	1·0 (ref)
Per 50 cells per µL[b]		0·86 (0·85–0·88)		0·87 (0·86–0·89)		0·90 (0·88–0·92)		0·91 (0·89–0·94)
Global p-value (AIC) 1[c]		<0·0001 (6625)		<0·0001 (6657)		<0·0001 (6708)		<0·0001 (6736)
Global p-value (AIC) 2[d]		<0·0001 (3495)		<0·0001 (3503)		<0·0001 (3519)		<0·0001 (3521)
CD4 count, cells per µL				**1620 d lag**		**1980 d lag**		**2340 d lag**
<50			18	3·5 (2·1–5·9)	11	2·3 (1·2–4·4)	6	1·4 (0·6–3·3)
50–<100			33	4·8 (3·2–7·3)	20	3·2 (1·9–5·3)	14	2·5 (1·4–4·5)
100–<200			66	3·1 (2·2–4·3)	49	2·4 (1·7–3·5)	35	2·0 (1·3–3·0)
200–<350			75	1·6 (1·2–2·3)	72	1·7 (1·2–2·4)	62	1·7 (1·2–2·4)
350–<500			64	1·4 (1·0–1·9)	49	1·1 (0·8–1·6)	41	1·1 (0·7–1·6)
≥500			76	1·0 (ref)	69	1·0 (ref)	59	1·0 (ref)
Per 50 cells per µL[b]				0·92 (0·90–0·95)		0·93 (0·91–0·96)		0·94 (0·91–0·97)
Global p-value (AIC) 1[c]				--		--		--
Global p-value (AIC) 2[d]				<0·0001 (3537)		<0·0001 (3549)		0·0022 (3563)
HIV RNA level, copies per mL		**180 d lag**		**540 d lag**		**900 d lag**		**1260 d lag**
≤500	297	1·0 (ref)	199	1·0 (ref)	155	1·0 (ref)	138	1·0 (ref)
>500–<10,000	102	1·3 (1·0–1·7)	109	2·0 (1·6–2·5)	112	2·4 (1·9–3·1)	84	1·8 (1·4–2·4)
10,000–<100,000	172	2·8 (2·3–3·4)	167	4·0 (3·2–4·9)	152	4·3 (3·4–5·4)	121	3·5 (2·7–4·5)
≥100,000	141	6·7 (5·4–8·3)	122	9·1 (7·2–11·6)	81	7·5 (5·7–9·9)	60	5·9 (4·3–8·1)
Per log10 copies per mL[b]		1·71 (1·61–1·82)		1·91 (1·78–2·04)		1·83 (1·69–1·97)		1·75 (1·61–1·90)
Global p-value (AIC) 1[c]		<0·0001 (6613)		<0·0001 (6582)		<0·0001 (6633)		<0·0001 (6684)
Global p-value (AIC) 2[d]		<0·0001 (3469)		<0·0001 (3451)		<0·0001 (3474)		<0·0001 (3503)
HIV RNA level, copies per mL				**1620 d lag**		**1980 d lag**		**2340 d lag**
≤500			116	1·0 (ref)	94	1·0 (ref)	75	1·0 (ref)
>500–<10,000			74	1·7 (1·3–2·4)	57	1·5 (1·1–2·1)	51	1·6 (1·1–2·3)
10,000–<100,000			105	3·4 (2·6–4·5)	92	3·5 (2·6–4·7)	76	3·4 (2·5–4·8)
≥100,000			37	4·3 (2·9–6·3)	27	3·7 (2·4–5·7)	15	2·4 (1·4–4·2)
Per log10 copies per mL[b]				1·65 (1·50–1·81)		1·59 (1·43–1·77)		1·51 (1·34–1·70)
Global p-value (AIC) 1[c]				--		--		--
Global p-value (AIC) 2[d]				<0·0001 (3523)		<0·0001 (3516)		<0·0001 (3525)

AIC=Akaike's information criterion. HR=hazard ratio. 95% CI=95% confidence interval. NHL=non-Hodgkin lymphoma.

[a] Each measure was individually included in a cohort-stratified Cox model adjusted for sex, race/ethnicity, and baseline age and calendar period. The N and number of NHL cases used for the model of each CD4 count and HIV RNA level measure was: 180 d lag (N= 102,131; number of NHL cases= 712), 540 d lag (N= 93,917; number of NHL cases= 597), 900 d lag (N= 79,458; number of NHL cases= 467), 1260 d lag (N= 68,585; number of NHL cases= 403), 1620 d lag (N= 59,696; number of NHL cases= 332), 1980 d lag (N= 51,326; number of NHL cases= 270), and 2340 d lag (N= 45,108; number of NHL cases= 217).

[b] P-trend<0·0001 for all measures.

[c] Global p-value and AIC 1 from models among persons with follow-up >1260 days (N= 68,585; number of NHL cases= 403).

[d] Global p-value and AIC 2 from models among persons with follow-up >2340 days (N= 45,108; number of NHL cases= 217).

Appendix Table 4. Cumulative and nadir/peak CD4 count and HIV RNA level measures during moving time windows, and overall NHL risk, NA-ACCORD, 1996-2014.

Measures	Early moving window[a]		Late moving window[a]		Overall moving window[a]	
	NHL cases	HR (95% CI)	NHL cases	HR (95% CI)	NHL cases	HR (95% CI)
CD4 count average, cells per µL						
<50	2	1·1 (0·3–4·5)	33	9·8 (6·5–14·7)	4	3·9 (1·4–10·9)
50–<100	19	5·0 (2·9–8·4)	37	5·9 (4·0–8·7)	16	5·2 (3·0–9·1)
100–<200	37	2·7 (1·8–4·1)	81	3·9 (2·9–5·3)	43	3·7 (2·5–5·6)
200–<350	59	1·8 (1·3–2·7)	85	1·8 (1·3–2·4)	57	2·0 (1·4–3·0)
350–<500	42	1·2 (0·8–1·7)	73	1·3 (1·0–1·8)	40	1·2 (0·8–1·8)
≥500	58	1·0 (ref)	94	1·0 (ref)	57	1·0 (ref)
Per 50 cells per µL[b]		0·92 (0·89–0·95)		0·87 (0·85–0·90)		0·90 (0·87–0·93)
Global p-value (AIC) 1[c]		--		<0·0001 (6663)		--
Global p-value (AIC) 2[d]		<0·0001 (3537)		<0·0001 (3506)		<0·0001 (3520)
Proportion of time CD4 count <200, cells per µL						
0·00	118	1·0 (ref)	184	1·0 (ref)	92	1·0 (ref)
>0·00–0·25	23	1·3 (0·8–2·0)	40	1·5 (1·1–2·1)	37	1·5 (1·0–2·2)
>0·25–0·50	16	1·9 (1·1–3·2)	27	2·2 (1·4–3·3)	22	2·3 (1·5–3·7)
>0·50–0·75	14	2·1 (1·2–3·7)	28	2·8 (1·9–4·2)	18	2·6 (1·6–4·3)
>0·75–<1·00	16	2·6 (1·5–4·4)	29	3·1 (2·1–4·7)	30	4·9 (3·2–7·4)
1·00	30	3·1 (2·0–4·6)	95	5·9 (4·6–7·6)	18	4·0 (2·4–6·6)
Per 20% of time CD4 <200 cells per µL[b]		1·25 (1·16–1·33)		1·39 (1·32–1·45)		1·37 (1·27–1·46)
Global p-value (AIC) 1[c]		--		<0·0001 (6664)		--
Global p-value (AIC) 2[d]		<0·0001 (3546)		<0·0001 (3492)		<0·0001 (3516)
CD4 count nadir, cells per µL						
<50	19	2·0 (1·1–3·7)	90	7·6 (5·2–11·3)	45	4·7 (2·6–8·4)
50–<100	37	4·3 (2·6–7·0)	54	5·2 (3·4–8·0)	38	4·9 (2·7–9·0)
100–<200	43	1·8 (1·1–2·9)	75	2·7 (1·8–4·1)	42	2·1 (1·2–3·9)
200–<350	56	1·3 (0·8–2·1)	84	1·7 (1·1–2·5)	44	1·4 (0·8–2·5)
350–<500	34	1·1 (0·7–1·8)	63	1·7 (1·1–2·5)	33	1·7 (0·9–3·1)
≥500	28	1·0 (ref)	37	1·0 (ref)	15	1·0 (ref)
Per 50 cells per µL[b]		0·91 (0·87–0·94)		0·85 (0·83–0·88)		0·86 (0·83–0·90)
Global p-value (AIC) 1[c]		--		<0·0001 (6673)		--
Global p-value (AIC) 2[d]		<0·0001 (3541)		<0·0001 (3490)		<0·0001 (3520)
HIV RNA level average, copies per mL						
≤500	59	1·0 (ref)	81	1·0 (ref)	41	1·0 (ref)
>500–<10,000	48	1·3 (0·9–1·9)	81	2·2 (1·6–3·0)	48	1·6 (1·1–2·5)
10,000–<100,000	87	3·1 (2·2–4·4)	153	5·4 (4·1–7·2)	100	4·7 (3·2–6·9)
≥100,000	23	4·8 (2·9–7·9)	88	16·6 (12·0–22·9)	28	9·9 (6·0–16·5)
Per log10 copies per mL[b]		1·73 (1·51–1·97)		2·36 (2·14–2·60)		2·20 (1·90–2·55)
Global p-value (AIC) 1[c]		--		<0·0001 (6522)		--
Global p-value (AIC) 2[d]		<0·0001 (3514)		<0·0001 (3416)		<0·0001 (3466)
Proportion of time HIV RNA level >500, copies per mL						
0·00	43	1·0 (ref)	65	1·0 (ref)	27	1·0 (ref)
>0·00–0·25	26	1·4 (0·8–2·3)	31	1·4 (0·9–2·2)	29	1·5 (0·9–2·5)
>0·25–0·50	14	1·1 (0·6–2·0)	39	2·8 (1·8–4·1)	25	2·1 (1·2–3·7)
>0·50–0·75	20	1·6 (0·9–2·8)	53	4·0 (2·7–5·8)	32	3·1 (1·8–5·2)
>0·75–<1·00	36	2·6 (1·6–4·1)	70	5·0 (3·5–7·2)	61	5·8 (3·6–9·4)
1·00	78	4·2 (2·8–6·3)	145	7·9 (5·8–10·9)	43	8·3 (5·0–14·0)
Per 20% of time HIV RNA >500 copies per mL[b]		1·31 (1·22–1·41)		1·48 (1·40–1·57)		1·50 (1·38–1·62)
Global p-value (AIC) 1[c]		--		<0·0001 (6619)		--
Global p-value (AIC) 2[d]		<0·0001 (3517)		<0·0001 (3470)		<0·0001 (3482)
HIV RNA level peak, copies per mL						
≤500	43	1·0 (ref)	65	1·0 (ref)	27	1·0 (ref)
>500–<10,000	27	1·0 (0·6–1·6)	46	1·6 (1·1–2·4)	21	1·2 (0·7–2·1)
10,000–<100,000	82	2·5 (1·7–3·7)	115	3·3 (2·4–4·5)	66	2·5 (1·6–4·1)
≥100,000	65	3·4 (2·3–5·1)	177	8·1 (6·0–11·0)	103	5·3 (3·4–8·4)
Per log10 copies per mL[b]		1·51 (1·34–1·70)		1·99 (1·82–2·17)		1·78 (1·57–2·03)
Global p-value (AIC) 1[c]		--		<0·0001 (6596)		--
Global p-value (AIC) 2[d]		<0·0001 (3523)		<0·0001 (3460)		<0·0001 (3493)

AIC=Akaike's information criterion. HR=hazard ratio. 95% CI=95% confidence interval. NHL=non-Hodgkin lymphoma.

[a] Each measure was individually included in a cohort-stratified Cox model adjusted for sex, race/ethnicity, and baseline age and calendar period. The N, number of NHL cases, and covered days range for each moving window were: early moving window (N= 45,108; number of NHL cases= 217; from 2340 to 1260 days in the past), late

moving window (N= 68,585; number of NHL cases= 403; from 1260 to 180 days in the past), and overall moving window (N= 45,108; number of NHL cases= 217; from 2340 to 180 days in the past).
[b] P-trend<0·0001 for all measures.
[c] Global p-value and AIC 1 from models among persons with follow-up >1260 days (N= 68,585; number of NHL cases= 403).
[d] Global p-value and AIC 2 from models among persons with follow-up >2340 days (N= 45,108; number of NHL cases= 217).

Appendix Table 5. Global p-values and Akaike's Information Criteria for models with two CD4 count and HIV RNA level measures, NA-ACCORD, 1996-2014.

Measures	Among persons with follow-up >1260 days[a]	
	Global p-value[b]	AIC[b]
CD4 count		
Single measure in model		
CD4 180 d lag	<0·0001	6625
Late CD4 average	<0·0001	6663
Late proportion of time CD4 <200 cells per μL	<0·0001	6664
Late CD4 nadir	<0·0001	6673
Two measures in model		
CD4 180 d lag +	<0·0001	6628
Late CD4 nadir	0·22	
CD4 180 d lag +	<0·0001	6629
Late proportion of time CD4 <200 cells per μL	0·30	
CD4 180 d lag +	<0·0001	6631
Late CD4 average	0·58	
HIV RNA level		
Single measure in model		
Late HIV RNA average	<0·0001	6522
HIV RNA 540 d lag	<0·0001	6582
Late HIV RNA peak	<0·0001	6596
Late proportion of time HIV RNA >500 copies per mL	<0·0001	6619
Two measures in model		
Late HIV RNA average +	<0·0001	6513
HIV RNA 540 d lag	0·0017	
Late HIV RNA average +	<0·0001	6518
Late proportion of time HIV RNA >500 copies per mL	0·019	
Late HIV RNA average +	<0·0001	6527
Late HIV RNA peak	0·75	

AIC=Akaike's information criterion.

[a] Late cumulative and nadir/peak measures were calculated during the moving window in the more recent past (i.e., from 1260 to 180 days in the past).

[b] From cohort-stratified Cox model adjusted for sex, race/ethnicity, and baseline age and calendar period. N= 68,585; number of NHL cases= 403.

Appendix Table 6. Sensitivity analysis: final model for overall NHL risk adjusted for HIV risk group, smoking, and cumulative ART use, NA-ACCORD, 1996-2014.

Overall NHL predictors[a]	Final model[b]		HIV risk group		Smoking		Cumulative ART use[e]
			Multiple imputation[c]	Unimputed as a separate category[d]	Multiple imputation[c]	Unimputed as a separate category[d]	
	NHL cases	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Recent CD4 count (i.e., CD4 count 180 d lag), cells per µL							
<50	60	3·2 (2·2–4·7)	3·2 (2·2–4·7)	3·2 (2·2–4·7)	3·0 (2·0–4·4)	3·1 (2·1–4·5)	3·3 (2·2–4·8)
50–<100	41	2·5 (1·7–3·8)	2·5 (1·7–3·8)	2·5 (1·7–3·8)	2·5 (1·6–3·7)	2·5 (1·7–3·7)	2·6 (1·7–3·9)
100–<200	59	1·5 (1·1–2·2)	1·5 (1·1–2·2)	1·5 (1·1–2·2)	1·5 (1·1–2·2)	1·5 (1·1–2·2)	1·6 (1·1–2·2)
200–<350	78	1·2 (0·9–1·6)	1·2 (0·9–1·6)	1·2 (0·9–1·6)	1·2 (0·9–1·6)	1·2 (0·9–1·6)	1·2 (0·9–1·6)
350–<500	70	1·1 (0·8–1·6)	1·1 (0·8–1·6)	1·1 (0·8–1·6)	1·1 (0·8–1·6)	1·1 (0·8–1·6)	1·1 (0·8–1·6)
≥500	95	1·0 (ref)	1·0 (ref)	1·0 (ref)	1·0 (ref)	1·0 (ref)	1·0 (ref)
Late HIV RNA level average, copies per mL[f]							
≤500	81	1·0 (ref)	1·0 (ref)	1·0 (ref)	1·0 (ref)	1·0 (ref)	1·0 (ref)
>500–<10,000	81	2·1 (1·5–2·8)	2·1 (1·5–2·9)	2·1 (1·5–2·9)	2·1 (1·5–2·9)	2·1 (1·5–2·8)	2·1 (1·5–3·0)
10,000–<100,000	153	4·2 (3·1–5·7)	4·2 (3·1–5·8)	4·2 (3·1–5·8)	4·2 (3·1–5·7)	4·2 (3·1–5·7)	4·5 (3·1–6·3)
≥100,000	88	9·6 (6·5–14·0)	9·6 (6·6–14·1)	9·6 (6·6–14·1)	9·5 (6·5–13·9)	9·5 (6·5–13·9)	10·2 (6·7–15·4)

AIC=Akaike's information criterion. HR=hazard ratio. 95% CI=95% confidence interval. NHL=non-Hodgkin lymphoma.

[a] Among persons with follow-up >1260 days (N= 68,585; number of NHL cases= 403).

[b] From cohort-stratified Cox model with the key predictors as covariates and adjusted sex, race/ethnicity, and baseline age and calendar period.

[c] Adjusted as final model plus HIV risk group or smoking via multiple imputation (Appendix p 1).

[d] Adjusted as final model plus HIV risk group or smoking with unimputed unknowns treated as a separate category.

[e] Adjusted as final model plus proportion of time on antiretroviral therapy during the late moving window (i.e., from 1260 to 180 days in the past). AIC= 6490 compared to 6488 for the final model.

[f] HIV RNA average during the late moving window (i.e., from 1260 to 180 days in the past).

NA-ACCORD COLLABORATING COHORTS AND REPRESENTATIVES

AIDS Clinical Trials Group Longitudinal Linked Randomized Trials: Constance A. Benson and Ronald J. Bosch
AIDS Link to the IntraVenous Experience: Gregory D. Kirk
Fenway Health HIV Cohort: Kenneth H. Mayer and Chris Grasso
HAART Observational Medical Evaluation and Research: Robert S. Hogg, P. Richard Harrigan, Julio SG Montaner, Benita Yip, Julia Zhu, Kate Salters and Karyn Gabler
HIV Outpatient Study: Kate Buchacz and Jun Li
HIV Research Network: Kelly A. Gebo and Richard D. Moore
Johns Hopkins HIV Clinical Cohort: Richard D. Moore
John T. Carey Special Immunology Unit Patient Care and Research Database, Case Western Reserve University: Benigno Rodriguez
Kaiser Permanente Mid-Atlantic States: Michael A. Horberg
Kaiser Permanente Northern California: Michael J. Silverberg
Longitudinal Study of Ocular Complications of AIDS: Jennifer E. Thorne
Multicenter Hemophilia Cohort Study–II: Charles Rabkin
Multicenter AIDS Cohort Study: Joseph B. Margolick, Lisa P. Jacobson and Gypsyamber D'Souza
Montreal Chest Institute Immunodeficiency Service Cohort: Marina B. Klein
Ontario HIV Treatment Network Cohort Study: Abigail Kroch, Ann Burchell, Adrian Betts and Joanne Lindsay.
Retrovirus Research Center, Bayamon Puerto Rico: Robert F. Hunter-Mellado and Angel M. Mayor
Southern Alberta Clinic Cohort: M. John Gill
Study of the Consequences of the Protease Inhibitor Era: Steven G. Deeks and Jeffrey N. Martin
Study to Understand the Natural History of HIV/AIDS in the Era of Effective Therapy: Jun Li and John T. Brooks
University of Alabama at Birmingham 1917 Clinic Cohort: Michael S. Saag, Michael J. Mugavero and James Willig
University of California at San Diego: William C. Mathews
University of North Carolina at Chapel Hill HIV Clinic Cohort: Joseph J. Eron and Sonia Napravnik
University of Washington HIV Cohort: Mari M. Kitahata, Heidi M. Crane and Daniel R. Drozd
Vanderbilt Comprehensive Care Clinic HIV Cohort: Timothy R. Sterling, David Haas, Peter Rebeiro and Megan Turner
Veterans Aging Cohort Study: Amy C. Justice, Robert Dubrow, and David Fiellin
Women's Interagency HIV Study: Stephen J. Gange and Kathryn Anastos

NA-ACCORD STUDY ADMINISTRATION

Executive Committee: Richard D. Moore, Michael S. Saag, Stephen J. Gange, Mari M. Kitahata, Keri N. Althoff, Michael A. Horberg, Marina B. Klein, Rosemary G. McKaig and Aimee M. Freeman
Administrative Core: Richard D. Moore and Aimee M. Freeman
Data Management Core: Mari M. Kitahata, Stephen E. Van Rompaey, Heidi M. Crane, Daniel R. Drozd, Liz Morton, Justin McReynolds and William B. Lober
Epidemiology and Biostatistics Core: Stephen J. Gange, Keri N. Althoff, Jennifer S. Lee, Bin You, Brenna Hogan, Jinbing Zhang, Jerry Jing, Elizabeth Humes, and Sally Coburn