Association of CD4$^+$ cell count and HIV viral load with risk of non-AIDS-defining cancers

Yunqing Ma[a], Jiajia Zhang[a,b], Xueying Yang[b,c], Shujie Chen[a], Sharon Weissman[b,d], Bankole Olatosi[b,e], Anthony Alberg[a] and Xiaoming Li[b,c]

Objectives: HIV-induced immunodeficiency contributes to an increased risk of non-AIDS-defining cancers (NADC). This study aims to identify the most predictive viral load (VL) or CD4$^+$ measures of NADC risk among people with HIV (PWH).

Design: Extracted from South Carolina electronic HIV reporting system, we studied adult PWH who were cancer-free at baseline and had at least 6 months of follow-up since HIV diagnosis between January 2005 and December 2020.

Methods: Using multiple proportional hazards models, risk of NADC was investigated in relation to 12 measures of VL and CD4$^+$ cell count at three different time intervals before NADC diagnosis. The best VL/CD4$^+$ predictor(s) and final model were determined using Akaike's information criterion.

Results: Among 10 413 eligible PWH, 449 (4.31%) developed at least one type of NADC. After adjusting for potential confounders, the best predictors of NADC were the proportion of days with viral suppression (hazard ratio [HR]: 0.47 (>25% and ≤50% vs. 0), 95% confidence interval [CI]: [0.28, 0.79]) and proportion of days with low CD4$^+$ cell count (AIC = 7201.35) (HR: 12.28 (>75% vs. = 0), 95% CI: [9.29, 16.23]).

Conclusions: VL and CD4$^+$ measures are strongly associated with risk of NADC. In analyses examining three time windows, proportion of days with low CD4$^+$ cell count was the best CD4$^+$ predictor for each time window. However, the best VL predictor varied across time windows. Thus, using the best combination of VL and CD4$^+$ measures for a specific time window should be considered when predicting NADC risk.

Keywords: CD4$^+$ cell count, HIV/AIDS, non-AIDS-defining cancers, South Carolina, viral load

Introduction

AIDS–related morbidity and mortality in people with HIV (PWH) have sharply declined since the introduction of antiretroviral therapy (ART) in 1996 [1–4]. By 2017, in the United States more than one–half of PWH in the United States (US) are 50 years of age or older [5]. The predicted proportion of older PWH (≥50) will increase from 28% in 2010 to 73% by 2030 [6]. With the improvement of life expectancy, non–AIDS-defining cancers (NADCs) have become increasingly common among PWH [7–9], and the overall relative risk for all NADC among PWH is ∼2-fold higher than in the general population. Shiels and colleagues predicted that the total proportion of NADC burden (NADC out of total cancer cases) in PWH will increase from 66.5% (NADC out of total cancer) cases in 2010 to 89.4% in 2030 [10], underscoring the importance of

[a]Department of Epidemiology and Biostatistics, [b]South Carolina SmatState Center for Healthcare Quality, [c]Department of Health Promotion, Education and Behavior, [d]Department of Internal Medicine, School of Medicine, and [e]Department of Health Services Policy and Management, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, USA.

Correspondence to Yunqing Ma, MS, Suite 527, 915 Greene Street, Columbia, SC 29208, USA.

E-mail: yunqing@email.sc.edu

NADC among PWH as a clinical and population health concern [11].

PWH are clearly at increased risk for NADC, and further research is warranted to more precisely pinpoint the relative contribution of long-term exposure to antiretroviral agents and HIV infection, which is characterized by immunosuppression and chronic immune activation [12−15]. Immune dysregulation is considered an important contributor to carcinogenesis [16]. According to a study on US veterans [17], HIV viral suppression was associated with reduced NADC. HIV treatment response is well defined in clinical practice using both viral load (VL) and $CD4^+$ cell count [18]. VL, measured in various ways including peak VL [19], recent VL [20], and cumulative VL [21,22], had been found related to cancer in different ways. Higher viral loads [23] and decreased duration of viral suppression [24] have been observed to be associated with increased cancer risk overall; and cumulative and early viral suppression [25] were also reported to be associated with low risk of specific cancers [14,22,26]. In early studies, lower $CD4^+$ cell count at the time of cancer diagnosis was not predictive of NADC [27,28]. In more recent studies, however, some time-updated $CD4^+$ cell count measures were independently associated with NADC risk with recent $CD4^+$ cell count being the most common $CD4^+$ cell count predictor for some specific NADCs [14,22,29]. Lower nadir $CD4^+$ cell count [20,30−32], cumulative exposure to low $CD4^+$ cell count (<200 cells/μl), and recent $CD4^+$ cell count [33] were found to be associated with an increased NADC risk. In addition, smoking, unhealthy alcohol use, opioid use, and depression are common among PWH and individually contribute to increased risk for non–AIDS-related comorbidities [34].

Many studies have already examined the influence of VL and $CD4^+$ cell count on NADC risk, but several key questions remain unanswered. When investigating the relationship of $CD4^+$ cell count and VL with NADC risk, cross-sectional studies are unable to capture the longitudinal measure of $CD4^+$/VL comprehensively (e.g. cumulative measures over a time course). Only few studies have examined recent and cumulative VL in relation to overall NADC risk independent of $CD4^+$ cell count [14,20]. Some previous studies have analyzed the relationship using both VL and $CD4^+$ measures but only at baseline [30]; there is absence of recent evidence. How different combination could impact the NADC risk prediction is unclear. A low $CD4^+$ cell count from a later time course was associated with an increased risk of NADC mortality [7]. However, little is known about the potential clinical value of using VL or $CD4^+$ cell count-related measures over a short period of time ahead of the NADC diagnosis.

NADC encompasses various cancers with their own characteristics [35], including those linked to HIV-related immunodeficiency, e.g. Hodgkin lymphoma, noncervical HPV-related cancers, and conjunctival cancers [36]. Some cancers are more prevalent due to aging rather than HIV and immunodeficiency. Certain cancers offer potential for early detection through targeted screening or early diagnosis methods, while others do not [37,38]. Grouping NADC in a detailed and specific manner is beneficial. Classification can be based on virus-caused cancers such as HPV and HBV [39].

Understanding the relationship between $CD4^+$, VL, and NADC is crucial for early detection of NADC risk and improving outcomes for PWH. This study, using 15 years of South Carolina statewide electronic health record (EHR) data, aims to identify key predictors of NADC risk based on $CD4^+$ cell count and VL measurements. By assessing various measures among PWH in South Carolina (such as baseline and cumulative VL and $CD4^+$ cell count, and duration of low $CD4^+$ or viral suppression), the study contributes evidence towards developing a clinical prediction model using $CD4^+$ and/or VL measures as predictors of NADC risk.

Methods

Study cohort
The population-based cohort was extracted through the integrated system of statewide electronic health record (EHR) data in SC. Beginning in 1986, SC Department of Health and Environmental Control's (SC DHEC) enhanced HIV/AIDS reporting system (eHARS) has collected a statewide confidential name-based reporting of HIV/AIDS, with $CD4^+$ cell count and viral load tests becoming required by the CDC since January 1, 2004 [40]. The de-identified EHR data from SC DHEC's eHARS and all payers' claim data were linked by SC Revenue and Fiscal Affairs Office (SC RFA). Details of data sources and data linkage are described elsewhere [41,42].

For this study, we included all adult PWH (age $\geq$ 18 at HIV diagnosis) who (1) were cancer-free at the beginning of follow-up, (2) had at least 6 months follow-up since HIV diagnosis, (3) had HIV diagnosed in SC from January 1, 2005 to December 31, 2020, and (4) with at least two VL and $CD4^+$ test records after HIV diagnosis but before NADC diagnosis or end of study, whichever occurred first. In total, there were 10 413 PWH who met the inclusion criteria.

Measures
Outcome
NADC was defined as cancers that do not belong to ADCs (i.e. non-Hodgkin lymphoma, Burkitt lymphoma, invasive cervical cancer, and Kaposi sarcoma) [43]. We identified 30 NADCs [44] using International Classification of Diseases, 9th and 10th revision

(ICD-9/10) Diagnostic Codes [45] (Table 1, Supplemental Digital Content, http://links.lww.com/QAD/C914). The entire follow-up time window of patients started from their initial HIV diagnosis date until the NADC diagnosis, death, or the end of study (December 31, 2020), whichever occurred first. Time to NADC was defined as number of days from initial HIV diagnosis to the first NADC diagnosis. Censoring also occurred due to death or the end of follow-up.

The categorized of NADCs comprises a heterogeneous group of malignancies. To attempt to achieve more refined resolution of the overall results, we also carried out separate analyses for virus-related NADCs (Table 8, Supplemental Digital Content, http://links.lww.com/QAD/C914) and for lung cancer (Table 9, Supplemental Digital Content, http://links.lww.com/QAD/C914). Virus-related NADCs were defined according to IARC Monograph 100b (Table 7, Supplemental Digital Content, http://links.lww.com/QAD/C914) [39].

CD4$^+$ and viral load measures

The historical CD4$^+$ and VL measures were defined within specific clinically relevant time windows before NADC diagnosis. We defined six measures for VL (i.e. baseline, recent, mean, and cumulative VL, proportion of days with viral suppression, and max VL) and six measures for CD4$^+$ cell count (i.e. baseline, recent, mean, and cumulative CD4$^+$ cell count, proportion of days with low CD4$^+$, and nadir CD4$^+$ cell count). Specifically, we focused on the historical CD4$^+$ and VL measures falling into three time windows: (1) entire time interval from HIV diagnosis to patient endpoint (NADC diagnosis, death, or end of study), (2) 6 months (180 days) before NADC diagnosis, and (3) 18 months (540 days) before NADC diagnosis.

During the specific time window, cumulative measures were defined as a weighted average calculated using the average of product of time and the value of measure. All VL measures were then categorized into four groups: <500 copies/ml, 500−10 000 copies/ml, 10 000−100 000 copies/ml, and ≥100 000 copies/ml; all CD4$^+$ measures were categorized into four groups: <200 cells/μl, 200 to 350 cells/μl, 350−500 cells/μl, and ≥ 500 cells/μl. We also constructed the proportion of days with low CD4$^+$ or viral suppression over the total follow-up days, which was categorized as 0, >0 and ≤25%, >25% and ≤50%, >50% and ≤75%, and >75% and ≤100%.

Other predictors

Other potential predictors included gender, age at HIV diagnosis, race, transmission mode for HIV, and residence area (urban vs. rural). In terms of substance use, we included alcohol use, tobacco use, and illicit drug abuse. Preexisting conditions, that is, hypothyroidism, hypertension, arthritis, chronic obstructive pulmonary disease (COPD), cardiovascular disease, renal disease,

diabetes mellitus, obesity, cerebrovascular disease, dyslipidemia, hepatitis C, and hepatitis B were identified via ICD-9/10 code and included in the analysis.

Statistical analysis

Distributions of demographic characteristics and different VL and CD4$^+$ measures between PWH with and without NADC (and the virus-related NADC subgroup along with lung cancer subgroup) were summarized and examined using analysis of variance (ANOVA). The associations between different measures of VL and CD4$^+$ cell count and NADC (and the virus-related NADC subgroup along with lung cancer subgroup) were examined using the proportional hazards (PH) model. The PH assumptions were validated through the Schoenfeld residuals [46]. The Wald test was used to calculate *P*-values for the association of each measure with NADC risk. One-sided log–rank test was used to calculate *P*-values for trend for each of the comparisons that showed evidence of a dose-response trend. In each model, all other potential predictors mentioned above were adjusted.
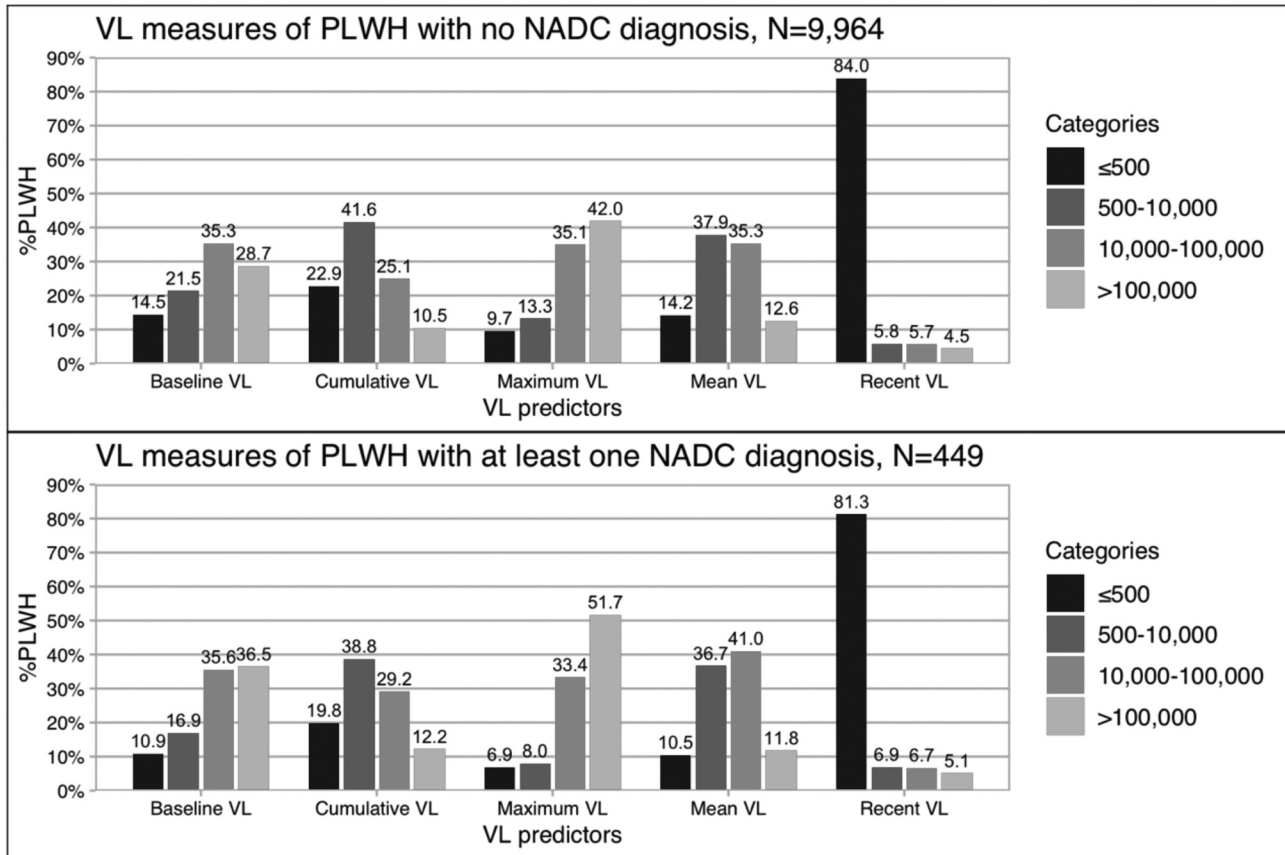
We consider each of the three time windows separately. In each time window defined above, we applied the following steps to choose the best combinations of CD4$^+$/VL predictors for NADC. First, each of six VL or six CD4$^+$ measures was modeled with NADC with PH model separately adjusting all potential confounders. Second, the significant measures of VL/CD4$^+$ cell count were selected and modeled via the PH model to determine the most significant measure. Then, we determine the best VL/CD4$^+$ measure based on Akaike's information criterion (AIC). A smaller AIC indicates a better model; a difference in AIC of >10 between two models is considered meaningful [47]. Last, we combine the selected best VL and CD4$^+$ measure and modeled in a final PH model adjusting all potential confounders. We used R version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria) for analyses. A two-sided *P*-value of 0.05 was employed to determine statistical significance.

Results

In this cohort of PWH, the majority were male (75.7%), Black American (69.5%), from urban areas (83.2%), and were men who have sex with men (MSM, 51.4%) (Table 1). The age at HIV diagnosis was between 18 and 30 years for 41.2% of the cohort. There were 449 (4.31%) individuals diagnosed with at least one type of NADC. There were 559 specific NADC diagnoses. The most common NADC diagnosis were lung and bronchus cancer (17.0%), Hodgkin lymphoma (8.1%), and breast cancer (7.3%) (Table 1, Supplemental Digital Content, http://links.lww.com/QAD/C914). PWH with NADC were older than those without NADC (40.1 vs. 34.9%, age > 50). PWH with NADC were less likely to be MSM (34.1 vs. 52.2%) compared with those without NADC.

Table 1. Social demographics and proportion of days with viral suppression/low CD4$^+$ cell count for the entire time window among eligible PWH in South Carolina.

Items	Total N (%)	NADC Yes	NADC No	P value*
N (%)	10 413 (100)	449 (4.3)	9964 (95.6)	
Sex				0.110
Male	7886 (75.7)	326 (72.6)	7560 (75.9)	
Female	2527 (24.3)	123 (27.4)	2404 (24.1)	
Age				<0.001
18–30	4287 (41.2)	53 (11.8)	4234 (42.5)	
30–40	2314 (22.2)	62 (13.8)	2252 (22.6)	
40–50	2107 (20.2)	150 (33.4)	1957 (19.6)	
50–60	1254 (12.0)	129 (28.3)	1125 (11.3)	
60+	451 (4.3)	55 (12.2)	396 (4.0)	
Race				0.003
White	2336 (22.4)	106 (23.6)	2230 (22.4)	
Black	7234 (69.5)	327 (72.8)	6907 (69.3)	
Hispanic	559 (5.4)	8 (1.8)	551 (5.5)	
Others	284 (2.7)	8 (1.8)	276 (2.8)	
Transmission mode				<0.001
Heterosexual	1983 (19.0)	132 (29.4)	1851 (18.6)	
MSM	5350 (51.4)	153 (34.1)	5197 (52.2)	
Injecting drug use/MSM	549 (5.3)	25 (5.6)	524 (5.3)	
Others	2531 (24.3)	139 (31.0)	2392 (24.0)	
Residence area				0.012
Rural	1753 (16.8)	95 (21.2)	1658 (16.6)	
Urban	8660 (83.2)	354 (78.8)	8306 (83.4)	
Proportion of days with viral suppression				<0.001
0	608 (5.8)	29 (6.5)	579 (5.8)	
0–25%	733 (7.0)	25 (5.6)	708 (7.1)	
25%–50%	1301 (12.5)	29 (6.5)	1272 (12.8)	
50%–75%	2117 (20.3)	45 (10.0)	2072 (20.8)	
>75%	5654 (54.3)	321 (71.5)	5333 (53.5)	
Proportion of days with low CD4$^+$ cell count				<0.001
0	5973 (57.4)	140 (31.2)	5833 (58.6)	
0–25%	2345 (22.5)	111 (24.7)	2234 (22.4)	
25%–50%	887 (8.5)	49 (10.9)	838 (8.4)	
50%–75%	577 (5.5)	39 (8.7)	538 (5.4)	
>75%	629 (6.0)	110 (24.5)	519 (5.2)	

*P values were calculated using ANOVA test (Pearson's chi-squared test).

PWH with and without NADC had similar patterns of each VL measure (Fig. 1). However, a higher proportion of individuals with max VL over 100 000 copies/ml was generally observed in PWH with NADC (e.g. maximum >100 000 copies/ml: 51.7 vs. 42.0%). The proportion of days with viral suppression (50–75%) for PWH with NADC (10.0%) was lower than those without NADC (20.8%). The patterns of CD4$^+$ measures for PWH with and without NADC had a slightly larger variation than VL measures (Fig. 2). For example, mean value of CD4$^+$ cell count was mostly between 200 and 350 cells/μl among PWH without NADC (51.9%), but <200 cells/μl for PWH with NADC (65.7%). The descriptive analysis results for 180 days' and 540 days' time windows were similar (Table 2, Supplemental Digital Content, http://links.lww.com/QAD/C914, Table 3, Supplemental Digital Content, http://links.lww.com/QAD/C914).

In the entire population, after adjusting for potential confounders, all 6 separate models for VL measures and all six models for separate CD4$^+$ measures showed significant associations with risk for any NADC (Table 2). All

predictors except for cumulative VL (P-value = 0.052) showed the significance of a dose–response trend. Proportion of days with viral suppression was the best VL measure among significant VL measures (AIC = 7378.94), and proportion of days with low CD4$^+$ cell count was the best CD4$^+$ cell count measure (AIC = 7201.35, Table 2). Thus, our final prediction model consisted of two predictors, proportion of days with viral suppression (>25 and ≤50% vs. 0; HR: 0.47, 95% CI: [0.28, 0.79]) and proportion of days with low CD4$^+$ cell count (>75% vs. 0; HR: 12.28, 95% CI: [9.29, 16.23]), adjusted for all potential confounders (Table 4, Supplemental Digital Content, http://links.lww.com/QAD/C914), which resulted in an AIC of 7209.34.

In the time window of 18 months prior to NADC diagnosis, the final model included baseline VL (>100 000 vs. ≤500; HR: 3.30, 95% CI: [2.42, 4.49], Table 3) and proportion of days with low CD4$^+$ cell count (>75% vs. 0; HR: 4.78, 95% CI: [3.49, 6.55]) with AIC = 6153.83. In the time window of 6 months prior to NADC diagnosis, the final model included cumulative VL (>10 000 and

Fig. 1. Percentage of PWH in different categories among five VL predictors for the entire time window. PWH, people with HIV; VL, vial load.

≤100 000 vs. ≤500; HR: 3.31, 95% CI: [2.37, 4.63], Table 3) and proportion of days with low $CD4^+$ cell count (>75% vs 0; HR: 5.30, 95% CI: [3.77, 7.43]) with AIC = 5071.07.

For the virus–related subgroup analysis, cumulative VL (>100 000 vs. ≤500; HR: 2.41, 95% CI: [1.21, 4.79], Table 8, Supplemental Digital Content, http://links.lww.com/QAD/C914) and proportion of days with low $CD4^+$ cell count (>75% vs. 0; HR: 22.74, 95% CI: [13.42, 38.51]) were included in the final model. For the lung cancer subgroup analysis, cumulative VL (>100 000 vs. ≤500; HR: 2.26, 95% CI: [1.11, 4.58], Table 9, Supplemental Digital Content, http://links.lww.com/QAD/C914) with and proportion of days with low $CD4^+$ cell count (>75% vs. 0; HR: 11.68, 95% CI: [6.53, 20.88]) were included in the final model.

Discussion

Based on population–based data from PWH in SC, we identified the proportion of days with viral suppression and low $CD4^+$ cell count as the top predictors for NADC risk. Ongoing viral replication and sustained low $CD4^+$

levels emerged as the most significant factors compared to other VL and $CD4^+$ measures.

In the analyses of either 6-month or 18-month time windows, proportion of days with low $CD4^+$ cell count was persistently the best $CD4^+$ cell count predictor in both models, implying more attention ought to be paid to the optimal immune recovery to protect against NADC development. This result reinforces prior findings from several studies [14,26]. Although the best VL predictor varied between 6-month (cumulative VL is the best) and 18-month (baseline VL is the best) time window, the slight difference of AICs when comparing the baseline VL model and cumulative VL model in either 6-month (i.e. 5151.65 vs. 5159.20) or 18-month (i.e. 6245.32 vs. 6246.00) time windows demonstrating the nearly equal importance of these two measures in predicting NADC risk. Similar results were observed in virus–related and lung cancer subgroups, indicating that even though NADCs are made up of malignancies with important differences in their etiologic pathways and clinical course, cumulative VL and proportion of days with low $CD4^+$ cell count retained consistent predictive ability in these two key subgroups.

The second key finding was that combining the best VL and $CD4^+$ cell count measures outperformed using VL or
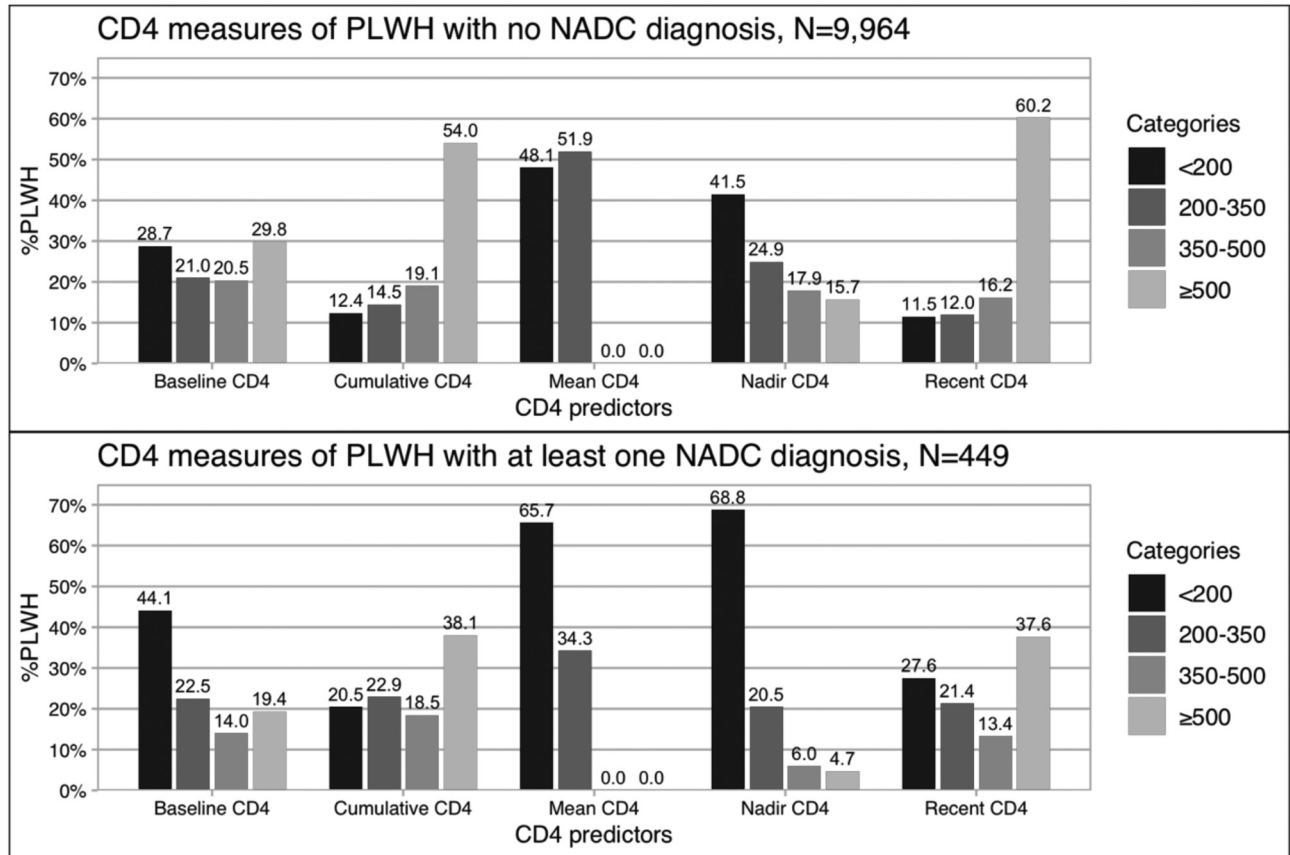
Fig. 2. Percentage of PWH in different categories among five CD4$^+$ predictors for the entire time window. PWH, people with HIV.

CD4$^+$ cell count alone for predicting NADC risk. Through a comprehensive assessment, the combination of VL and CD4$^+$ cell count (i.e. proportion of days with viral suppression/low CD4$^+$) emerged as the best predictors among significant variables, surpassing the predictive ability of VL or CD4$^+$ measures alone. This supports previous observations that long-term VL status and sustained CD4$^+$ cell count are both important factors in predicting NADC risk [14,25,26].

The innovative approach involved evaluating prediction models in different time windows relative to NADC diagnosis. Findings indicated that the best measure of CD4$^+$ remained consistent across all three time windows. Although the best measure of VL varied, the minimal changes of AICs in the separate models of baseline VL and cumulative VL highlighting the importance of both timely HIV diagnosis and ART initiation to avoid the rapid viral replication after HIV diagnosis and maintain an optimal viral suppression over time to reduce the risk of NADC.

Strengths of our study included its large size, coverage of almost two decades, validated HIV diagnosis, longitudinal assessment of multiple VL and CD4$^+$ measures, and population-based study. The innovative approach of

investigating time windows based on time to NADC diagnosis merits further evaluation as it could be superior to examining these measures during the entire time courses, which varies by individual.

Our study's limitations included unavailability of information on antiretroviral therapy (ART), although we assume that those with suppressed VL were taking ART. Other limitations included possible selection bias resulting from exclusion criteria, and potentially missing variable, such as smoking status, due to incomplete data. Additionally, our data only pertained to SC, leading to incomplete information for individuals who migrated in and out of the state. The small number of total NADC cases in the cohort prevented analyses of each specific type of cancer, but subgroup analyses that were possible indicated the overall findings were consistent for viral associated NADCs and for lung cancer. Even so, there were limitations to grouping the virus–related cancers due to lack of specificity in defining specific malignancies using large claims databases. It is important to consider that clinical cancer is the terminal result of a complex series of steps occurring during a preclinical 'latent period' and full consideration of the impact of VL and CD4$^+$ in relation to the time course of carcinogenesis warrants further investigation.

Table 2. Results for predictors selection with best VL/CD4+ cell count predictors of the entire time window.

VL predictors	HR (95% CI)[a,b,c]	AIC[c,d] (trend test P-values[f])	CD4+ predictors	HR (95% CI)[a,b,c]	AIC[c,d] (trend test P-values[f])
Baseline VL (viral copies/ml)		7429.63	Baseline CD4+ cell count (cells/μl)		7421.44
≤500	1	(<0.0001)	<200	1	(<0.0001)
500–10 000	1.06 (0.90, 1.53)		200–50	0.96 (0.75, 0.92)*	
10 000–100 000	1.25 (0.90, 1.73)		350–00	0.69 (0.51, 0.92)*	
>100 000	1.49 (1.09, 2.07)*		≥500	0.62 (0.48, 0.81)*	
Maximum VL (viral copies/ml)		7428.05	Nadir CD4+ cell count (cells/μl)		7371.20
≤500	1	(0.0004)	<200	1	(<0.0001)
500–10 000	0.93 (0.57, 1.52)		200–350	0.62 (0.49, 0.78)*	
10 000–100 000	1.25 (0.85, 1.85)		350–500	0.34 (0.23, 0.51)*	
>100 000	1.49 (1.02, 2.18)*		≥500	0.31 (0.20, 0.49)*	
Recent VL (viral copies/ml)		7427.15	Recent CD4+ cell count (cells/μl)		7315.74
≤500	1	(0.0374)	<200	1	(<0.0001)
500–10 000	1.50 (1.03, 2.18)*		200–350	0.63 (0.48, 0.82)*	
10 000–100 000	1.43 (0.98, 2.09)		350–500	0.33 (0.24, 0.45)*	
>100 000	1.78 (1.17, 2.73)*		≥500	0.27 (0.21, 0.34)*	
Mean VL (viral copies/ml)		7427.20	Mean CD4+ cell count (cells/μl)		7402.51
≤500	1	(0.0036)	<200	1	(<0.0001)
500–10 000	1.23 (0.89, 1.71)		200–350	0.56 (0.46, 0.68)*	
10 000–100 000	1.49 (1.08, 2.06)*				
>100 000	1.82 (1.22, 2.70)*				
Proportion of days with VL suppression		**7378.94***	**Proportion of days with low CD4+ cell count**		**7201.35***
0	1	(<0.0001)	0	1	(<0.0001)
0–0.25	0.61 (0.36, 1.04)		0–0.25	1.21 (0.94, 1.56)	
0.25–0.50	0.38 (0.23, 0.64)*		0.25–0.50	1.79 (1.29, 2.48)*	
0.50–0.75	0.30 (0.19, 0.49)*		0.50–0.75	2.49 (1.74, 3.57)*	
≥0.75	0.82 (0.56, 1.21)		≥0.75	9.46 (7.26, 12.34)*	
Cumulative VL (viral copies/ml)		7424.84	Cumulative CD4+ cell count (cells/μl)		7390.08
≤500	1	(0.0518)	<200	1	(<0.0001)
500–10 000	1.01 (0.78, 1.31)		200–350	0.74 (0.56, 0.64)*	
10 000–100 000	1.37 (1.04, 1.80)*		350–500	0.48 (0.35, 0.64)*	
>100 000	1.65 (1.17, 2.33)*		≥500	0.42 (0.32, 0.54)*	
Baseline model[e]		7433.10	Baseline model[e]		7433.10

VL, vial load.
[a]HR, hazard ratio, CI, confidence interval.
[b]*Significant predictors when α = 0.05.
[c]HRs and AIC were calculated through multiple PH models, that is, all significant predictors were modeled respectively adjusting all potential confounders.
[d]*The best VL/CD4+ cell count predictor based on AIC.
[e]Baseline model was a PH model using only confounders.
[f]P-values for trend test were calculated using one-sided log-rank test.

Table 3. Final models with best VL and CD4+ predictors for different time windows.

Entire time window, N = 10413		18 month time window, N = 8143		6 month time window, N = 6309	
Best predictors	HR (95% CI)[1]	Best predictors	HR (95% CI)[1]	Best predictors	HR (95% CI)[1]
Proportion of days with viral suppression		Baseline VL (viral copies/ml)		Cumulative VL (viral copies/ml)	
0	1	≤500	1	≤500	1
0–25%	0.54 (0.31, 0.92)*	500–10 000	2.37 (1.67, 3.37)*	500–10 000	2.81 (1.91, 4.15)*
25%–50%	0.47 (0.28, 0.79)*	10 000–100 000	3.01 (2.21, 4.10)*	10 000–100 000	3.31 (2.37, 4.63)*
50%–75%	0.52 (0.32, 0.83)*	>100 000	3.30 (2.42, 4.49)*	>100 000	2.15 (1.45, 3.18)*
>75%	1.71 (1.14, 2.56)*				
Proportion of days with low CD4+ cell count		Proportion of days with low CD4+ cell count		Proportion of days with low CD4+ cell count	
=0	1	=0	1	=0	1
0–25%	1.12 (0.87, 1.45)	0–25%	1.18 (0.90, 1.54)	0–25%	1.12 (0.84, 1.50)
25%–50%	2.30 (1.65, 3.21)*	25%-50%	1.42 (1.00, 2.01)*	25%-50%	1.51 (1.03, 2.20)*
50%–75%	3.95 (2.71, 5.76)*	50%-75%	1.74 (1.18, 2.56)*	50%-75%	1.90 (1.26, 2.87)*
>75%	12.28 (9.29, 16.23)*	>75%	4.78 (3.49, 6.55)*	>75%	5.30 (3.77, 7.43)*
Final model AIC	**7104.14**	**Final model AIC**	**6153.83**	**Final model AIC**	**5071.07**

CI, confidence interval; HR, hazard ratio; VL, viral load.

In summary, our findings indicate that considering both VL and $CD4^+$ cell count measures is preferable for NADC risk prediction compared to using either measure alone. Combining the best VL and $CD4^+$ cell count predictors enhances prediction performance compared to using VL or $CD4^+$ alone. The optimal measure of VL may vary depending on the timing of NADC diagnosis, making it important to prioritize long-term VL status and sustained $CD4^+$ cell count across different time windows. The evidence on this topic will become more clinically meaningful via future research that (1) moves past the broad categorization of NADCs to investigate either individual malignancies or meaningful groupings of malignancies and (2) has longer-term follow-up to evaluate these associations more completely throughout the carcinogenic process.

Acknowledgements

Conflicts of interest
There are no conflicts of interest.

References

1. Franceschi S, Lise M, Clifford GM, Rickenbach M, Levi F, Maspoli M, et al. **Changing patterns of cancer incidence in the early- and late-HAART periods: the Swiss HIV Cohort Study.** Br J Cancer 2010; **103**:416–422.
2. Park LS, Tate JP, Sigel K, Rimland D, Crothers K, Gibert C, et al. **Time trends in cancer incidence in persons living with HIV/AIDS in the antiretroviral therapy era: 1997–2012.** AIDS 2016; **30**:1795–1806.
3. Robbins HA, Shiels MS, Pfeiffer RM, Engels EA. **Epidemiologic contributions to recent cancer trends among HIV-infected people in the United States.** AIDS 2014; **28**:881–890.
4. Semeere AS, Busakhala N, Martin JN. **Impact of antiretroviral therapy on the incidence of Kaposi's sarcoma in resource-rich and resource-limited settings.** Curr Opin Oncol 2012; **24**:522–530.
5. CDC. US Centers for Disease Control and Prevention. HIV surveillance report, 2018 (updated). 2020. Available at: https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html. [Accessed 6 August 2022]
6. Smit M, Brinkman K, Geerlings S, Smit C, Thyagarajan K, Sighem A van, et al. **Future challenges for clinical care of an ageing population infected with HIV: a modelling study.** Lancet Infect Dis 2015; **15**:810–818.
7. Crum-Cianflone NF, Hullsiek KH, Marconi V, Weintrob A, Ganesan A, Barthel RV, et al. **Trends in the incidence of cancers among HIV-infected persons and the impact of antiretroviral therapy: authors' reply.** AIDS 2009; **23**:1791.
8. Engels EA, Biggar RJ, Hall HI, Cross H, Crutchfield A, Finch JL, et al. **Cancer risk in people infected with human immunodeficiency virus in the United States.** Int J Cancer 2008; **123**:187–194.
9. Silverberg MJ, Neuhaus J, Bower M, Gey D, Hatzakis A, Henry K, et al. **Risk of cancers during interrupted antiretroviral therapy in the SMART study.** AIDS 2007; **21**:1957–1963.
10. Shiels MS, Islam JY, Rosenberg PS, Hall HI, Jacobson E, Engels EA. **Projected cancer incidence rates and burden of incident cancer cases in HIV-infected adults in the united states through 2030.** Ann Intern Med 2018; **168**:866–873.
11. Horner M-J, Chasimpha S, Spoerri A, Edwards J, Bohlius J, Tweya H, et al. **High cancer burden among antiretroviral therapy users in Malawi: a record linkage study of observational human immunodeficiency virus cohorts and cancer registry data.** Clin Infect Dis 2019; **69**:829–835.
12. Chiao EY, Krown SE, Stier EA, Schrag D. **A Population-based analysis of temporal trends in the incidence of squamous anal canal cancer in relation to the HIV epidemic.** JAIDS J Acquir Immune Defic Syndr 2005; **40**:451–455.
13. Clifford GM, Polesel J, Rickenbach M, Dal Maso L, Keiser O, Kofler A, et al. **Cancer risk in the Swiss HIV cohort study: associations with immunodeficiency, smoking, and highly active antiretroviral therapy.** J Natl Cancer Inst 2005; **97**:425–432.
14. Guiguet M, Boué F, Cadranel J, Lang J-M, Rosenthal E, Costagliola D, et al. **Effect of immunodeficiency, HIV viral load, and antiretroviral therapy on the risk of individual malignancies (FHDH-ANRS CO4): a prospective cohort study.** Lancet Oncol 2009; **10**:1152–1159.
15. Long JL, Engels EA, Moore RD, Gebo KA. **Incidence and outcomes of malignancy in the HAART era in an urban cohort of HIV-infected individuals.** AIDS 2008; **22**:489–496.
16. Zanni MV, Grinspoon SK. **HIV-specific immune dysregulation and atherosclerosis.** Curr HIV/AIDS Rep 2012; **9**:200–205.
17. Chiao EY, Coghill A, Kizub D, Fink V, Ndlovu N, Mazul A, Sigel K. **The effect of non-AIDS-defining cancers on people living with HIV.** Lancet Oncol 2021; **22**:e240–e253.
18. Grabar S, Moing VL, Goujard C, Leport C, Kazatchkine MD, Costagliola D, et al. **Clinical outcome of patients with HIV-1 infection according to immunologic and virologic response after 6 months of highly active antiretroviral therapy.** Ann Intern Med 2000; **133**:401–410.
19. Achenbach CJ, Buchanan AL, Cole SR, Hou L, Mugavero MJ, Crane HM, et al. **HIV viremia and incidence of non-Hodgkin lymphoma in patients successfully treated with antiretroviral therapy.** Clin Infect Dis 2014; **58**:1599–1606.
20. Bruyand M, Thiébaut R, Lawson-Ayayi S, Joly P, Sasco A-J, Mercié P, et al. **Role of uncontrolled HIV RNA level and immunodeficiency in the occurrence of malignancy in HIV-infected patients during the combination antiretroviral therapy era: Agence Nationale de Recherche sur le Sida (ANRS) CO3 Aquitaine Cohort.** Clin Infect Dis 2009; **49**:1109–1116.
21. Kowalkowski M, Day R, Du X, Chan W, Chiao E. **Cumulative HIV viremia and non-AIDS-defining malignancies among a sample of HIV-infected male veterans.** J Acquir Immune Defic Syndr 2014; **67**:204–211.
22. Zoufaly A, Stellbrink H-J, Heiden MA der, Kollan C, Hoffmann C, van Lunzen J, et al. **Cumulative HIV viremia during highly active antiretroviral therapy is a strong predictor of AIDS-related lymphoma.** J Infect Dis 2009; **200**:79–87.
23. Reekie J, Gatell J, Yust I, Bakowska E, Rachmanova A, Losso M, et al. **Fatal and nonfatal AIDS and non-AIDS events in HIV-1 infected patients with high CD4 counts.** J Int AIDS Soc 2010; **13**:O39.
24. Reekie J, Gatell JM, Yust I, Bakowska E, Rakhmanova A, Losso M, et al. **Fatal and nonfatal AIDS and non-AIDS events in HIV-1-positive individuals with high CD4 cell counts according to viral load strata.** AIDS 2011; **25**:2259–2268.
25. Park LS, Tate JP, Sigel K, Brown ST, Crothers K, Gibert C, et al. **Association of viral suppression with lower AIDS-defining and non–AIDS-defining cancer incidence in HIV-infected veterans.** Ann Intern Med 2018; **169**:87–96.
26. Engels EA, Pfeiffer RM, Landgren O, Moore RD. **Immunologic and virologic predictors of AIDS-related non-Hodgkin lymphoma in the HAART era.** J Acquir Immune Defic Syndr 2010; **54**:78–84.

27. Burgi A, Brodine S, Wegner S, Milazzo M, Wallace MR, Spooner K, *et al.* **Incidence and risk factors for the occurrence of non-AIDS-defining cancers among human immunodeficiency virus-infected individuals.** *Cancer* 2005; **104**:1505–1511.
28. Mbulaiteye SM, Biggar RJ, Goedert JJ, Engels EA. **Immune deficiency and risk for malignancy among persons with AIDS.** *J Acquir Immune Defic Syndr* 2003; **32**:527–533.
29. Kirk O, Pedersen C, Cozzi-Lepri A, Antunes F, Miller V, Gatell JM, *et al.* **Non-Hodgkin lymphoma in HIV-infected patients in the era of highly active antiretroviral therapy.** *Blood* 2001; **98**:3406–3412.
30. Dauby N, De Wit S, Delforge M, Necsoi VC, Clumeck N. **Characteristics of non-AIDS-defining malignancies in the HAART era: a clinico-epidemiological study.** *J Int AIDS Soc* 2011; **14**:16.
31. Neuhaus J, Angus B, Kowalska JD, LA Rosa A, Sampson J, Wentworth D, *et al.* **Risk of all-cause mortality associated with nonfatal AIDS and serious non-AIDS events among adults infected with HIV.** *AIDS* 2010; **24**:697–706.
32. Powles T, Robinson D, Stebbing J, Shamash J, Nelson M, Gazzard B, *et al.* **Highly active antiretroviral therapy and the incidence of non-AIDS-defining cancers in people with HIV infection.** *HIV Med* 2009; **27**:884–890.
33. Krishnan S, Schouten JT, Jacobson DL, Benson CA, Collier AC, Koletar SL, *et al.* **Incidence of non-AIDS-defining cancer in antiretroviral treatment-naïve subjects after antiretroviral treatment initiation: an ACTG Longitudinal Linked Randomized Trials Analysis.** *Oncology* 2011; **80**:42–49.
34. Chichetto NE, Polanka BM, So-Armah KA, Sung M, Stewart JC, Koethe JR, *et al.* **Contribution of behavioral health factors to non-AIDS-related comorbidities: an updated review.** *Curr HIV/AIDS Rep* 2020; **17**:354–372.
35. Mitsuyasu RT. **Non-AIDS-defining malignancies in HIV.** *Top HIV Med* 2008; **16**:117–121.
36. Remick SC. **Non–AIDS-defining cancers.** *Hematol Oncol Clin North Am* 1996; **10**:1203–1213.
37. Ceccarelli M, Venanzi Rullo E, Marino MA, d'Aleo F, Pellicanò GF, D'Andrea F, *et al.* **Non-AIDS defining cancers: a comprehensive update on diagnosis and management.** *Eur Rev Med Pharmacol Sci* 2020; **24**:3849–3875.
38. Shiels MS, Althoff KN, Pfeiffer RM, Achenbach CJ, Abraham AG, Castilho J, *et al.* **HIV infection, immunosuppression, and age at diagnosis of non-AIDS-defining cancers.** *Clin Infect Dis* 2017; **64**:468–475.
39. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. **Biological agents. volume 100 B. a review of human carcinogens.** *IARC Monogr Eval Carcinog Risks Hum* 2012; **100**:1–441.
40. Centers for Disease Control and Prevention. **Missed opportunities for earlier diagnosis of HIV Infection - South Carolina, 1997–2005.** *Morb Mortality Weekly Rep* 2006; 1269–1272.
41. Olatosi B, Zhang J, Weissman S, Hu J, Haider MR, Li X. **Using big data analytics to improve HIV medical care utilisation in South Carolina: a study protocol.** *BMJ Open* 2019; **9**:e027688.
42. Zhang J, Olatosi B, Yang X, Weissman S, Li Z, Hu J, *et al.* **Studying patterns and predictors of HIV viral suppression using A Big Data approach: a research protocol.** *BMC Infect Dis* 2022; **22**:122.
43. NIH. Definition of non-AIDS-defining cancer – NCI dictionary of cancer terms – National Cancer Institute. 2011. Available at: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/nonaids-defining-cancer. [Accessed 15 April 2022]
44. Nkwonta CA, Zhang J, Chen S, Weissman S, Olatosi B, Li X. **Prevalence and trend of AIDS-defining cancers and non-AIDS-defining cancers and their association with antiretroviral therapy among people living with HIV in South Carolina: a population-based cohort study.** *AIDS Care* 2022; **35**:753–763.
45. Glasheen WP, Cordier T, Gumpina R, Haugh G, Davis J, Renda A. **Charlson comorbidity index: ICD-9 update and ICD-10 translation.** *Am Health Drug Benefits* 2019; **12**:188–197.
46. Grambsch PM, Therneau TM. **Proportional hazards tests and diagnostics based on weighted residuals.** *Biometrika* 1994; **81**:515–526.
47. Burnham KP, Anderson DR. **Multimodel inference: understanding AIC and BIC in model selection.** *Sociol Methods Res* 2004; **33**:261–304.