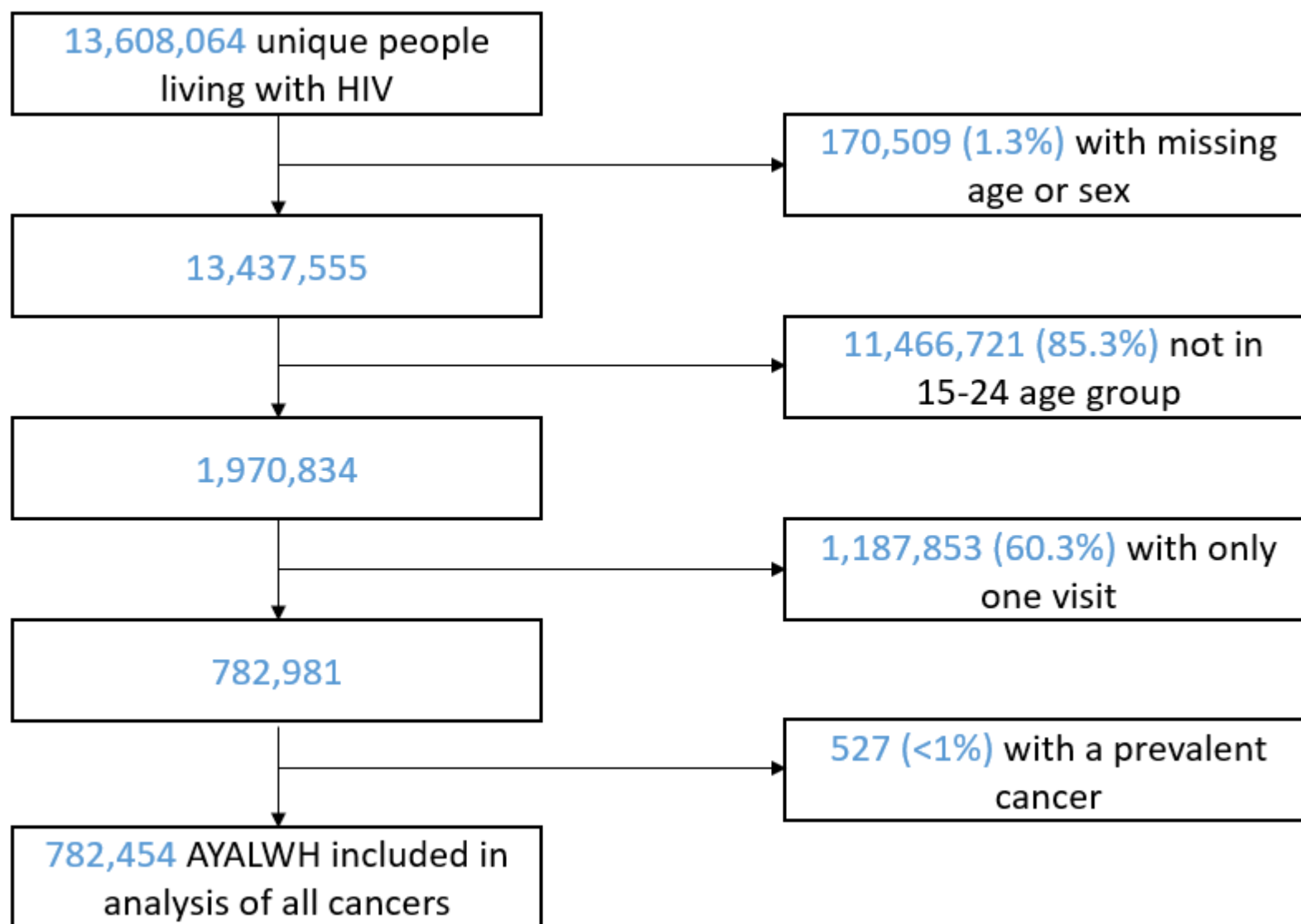**Figure:** Selection of adolescents and young adults living with HIV (AYALWH) in the South African HIV Cancer Match study who were eligible for inclusion in the overall cancer analysis.
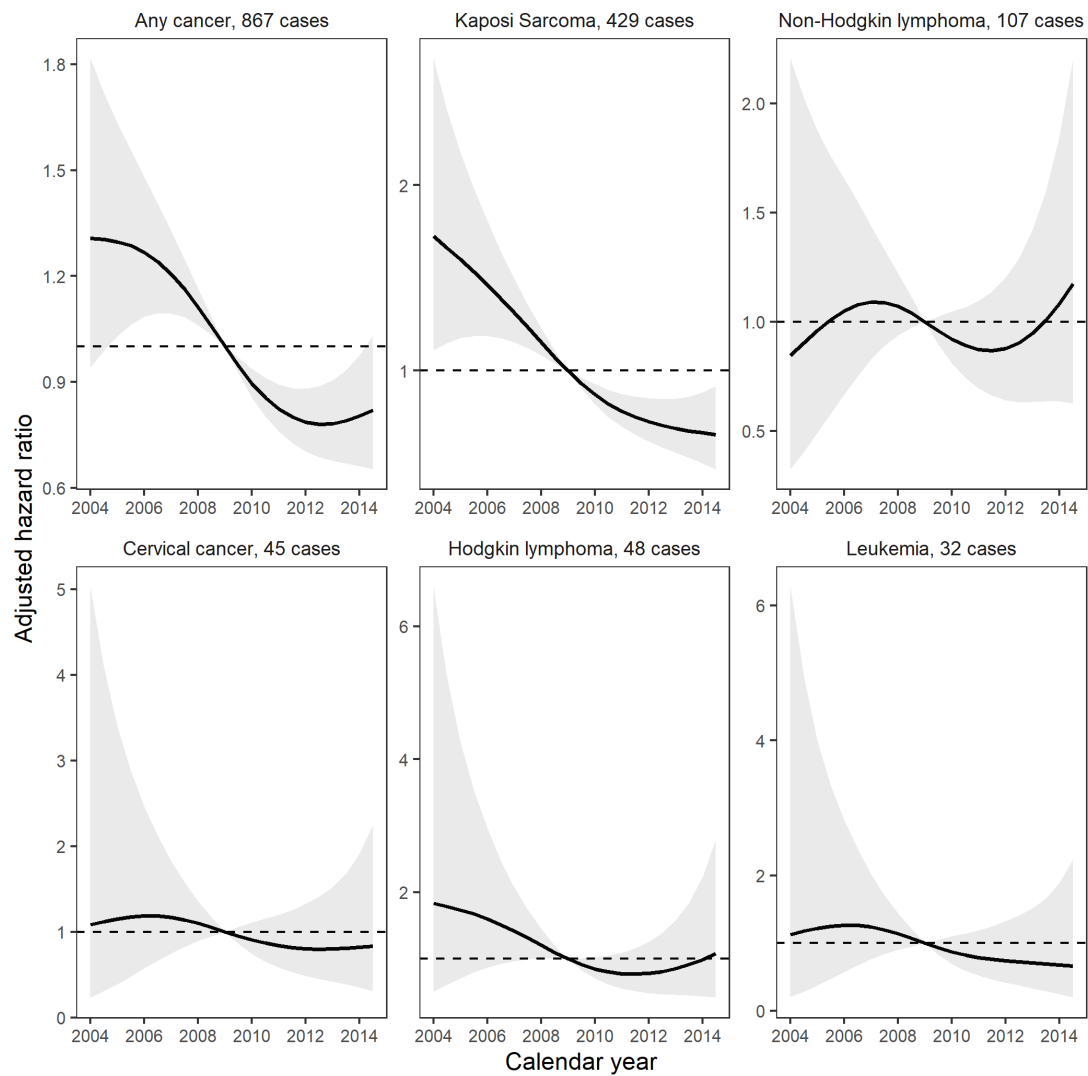
Table: Unadjusted hazard ratios for developing different cancer types among adolescents and young adults living with HIV.

	Hazard ratios (95% confidence intervals)					
	Any cancer	Kaposi sarcoma	Cervical cancer	Non-Hodgkin lymphoma	Hodgkin lymphoma	Leukaemia
Age [years]						
15-19	1	1	†	1	†	1
20-24	1.43 (1.17-1.75)	1.55 (1.16-2.08)		1.14 (0.69-1.90)		1.19 (0.46-3.12)
Sex						
Female	1	1	-	1	1	1
Male	2.38 (2.01-2.82)	2.35 (1.84-2.99)	-	3.54 (2.31-5.43)	5.03 (2.75-9.23)	5.90 (2.87-12.1)
Calendar period						
2004 – 2007	1	1	1	1	1	1
2008 – 2011	0.69 (0.59-0.80)	0.61 (0.49-0.75)	1.18 (0.58-2.38)	0.67 (0.43-1.05)	0.59 (0.30-1.14)	0.66 (0.31-1.41)
2012 – 2014	0.66 (0.54-0.81)	0.52 (0.39-0.70)	0.63 (0.21-1.86)	0.82 (0.48-1.39)	0.68 (0.30-1.55)	0.38 (0.12-1.20)
CD4 cell count [cells/µL]						
<200			1	1	1	1
200-499	†	†	0.42 (0.21-0.83)	0.42 (0.27-0.66)	0.41 (0.21-0.83)	0.67 (0.25-1.81)
≥500			0.19 (0.07-0.53)	0.26 (0.14-0.48)	0.44 (0.20-0.99)	1.23 (0.81-3.33)
Missing			0.56 (0.21-1.45)	0.68 (0.37-1.25)	0.61 (0.23-1.59)	1.19 (0.33-4.21)

Age was treated as a time-updated variable in all models.

† To preserve proportional hazards, we stratified the model for any cancer and for Kaposi sarcoma with respect to CD4 cell count and the model for cervical cancer and Hodgkin lymphoma with respect to age.

Figure: Hazard ratios (solid lines) with 95% confidence intervals (shaded areas) for the risk of developing cancer, comparing calendar years (time-updated) to the reference year of 2009. The models for any cancer, Kaposi sarcoma, non-Hodgkin lymphoma and Hodgkin lymphoma are adjusted for baseline age, sex, and baseline CD4 cell count. The model for cervical cancer is adjusted for age and baseline CD4 cell count. The model for leukaemia is not adjusted for other variables.

Figure: Cumulative hazard functions for the incidence of any cancer, stratified by baseline CD4 cell count category. The shaded areas represent 95% confidence intervals.