Received: 29 March 2023 | Revised: 12 July 2023 | Accepted: 31 July 2023

DOI: 10.1002/ijc.34712

RESEARCH ARTICLE

Cancer Epidemiology

IJC INTERNATIONAL JOURNAL of CANCER

Gynaecologic and breast cancers in women living with HIV in South Africa: A record linkage study

Tafadzwa G. Dhokotera[1,2,3,4] | Mazvita Muchengeti[4,5,6] | Maša Davidović[1,2,7] | Eliane Rohner[8] | Victor Olago[4,5] | Matthias Egger[8,9,10] | Julia Bohlius[1,2,8]

[1]Swiss Tropical and Public Health Institute, Allschwil, Switzerland

[2]University of Basel, Basel, Switzerland

[3]Graduate School for Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland

[4]National Cancer Registry, National Health Laboratory Service, Johannesburg, South Africa

[5]School of Public Health, University of the Witwatersrand, Johannesburg, South Africa

[6]South African DSI NRF Centre of Excellence in Epidemiological Modelling and Analysis, Stellenbosch University, Stellenbosch, South Africa

[7]Graduate School for Health Sciences, University of Bern, Bern, Switzerland

[8]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

[9]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

[10]Centre for Infectious Disease Epidemiology and Research, School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa

Correspondence
Tafadzwa G. Dhokotera, Swiss Tropical and Public Health Institute, Kreuzstrasse 2, 4123 Allschwil, Switzerland.
Email: tafadzwagladys.dhokotera@swisstph.ch

Funding information
CRDF Global, Grant/Award Number: HIV_DAA3-16-62705-1; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 801076; National Institutes of Health, Grant/Award Number: U01AI069924; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Numbers: 320030_169967, 207285; NIH Administrative Supplement, Grant/Award Number: U01AI069924-09

Abstract

Breast and gynaecologic cancers account for approximately half of all cancers diagnosed amongst women in South Africa, many of whom also live with HIV. We aimed to determine the incidence of and risk factors for developing breast and gynaecologic cancers in women living with HIV (WLHIV) in South Africa. This is a longitudinal analysis of the South African HIV Cancer Match study including women aged ≥15 years with two or more HIV-related laboratory tests. We used Cox proportional hazard models to determine the association of Human Papilloma Virus (HPV)-related and hormone-related gynaecologic cancer with patient- and municipal-level characteristics. From 3 447 908 women and 10.5 million years of follow-up, we identified 11 384 incident and 7612 prevalent gynaecologic and breast cancers. The overall crude incidence rate was 108/1 00 000 person-years (pyears) (95% confidence interval [CI]: 106-110), with the highest incidence observed for cervical cancer (70/1 00 000 pyears; 95% CI: 68.5-71.7). Low CD4 cell counts and high HIV RNA viral loads increased the risk of cervical and other HPV-related cancers. Age was associated with both HPV-related and hormone-related cancers. Women accessing health facilities in high socioeconomic position (SEP) municipalities were more likely

Abbreviations: ART, antiretroviral treatment; HPV, human papilloma virus; ICD-O-3, international classification of diseases for oncology version 3; NADC, non-AIDS defining cancer; NCR, National Cancer Registry; NHLS, National Health Laboratory Service; SAM, South African HIV cancer match study; SEP, socioeconomic position; WLHIV, women living with HIV.

to be diagnosed with HPV-related cancers and breast cancer than women accessing care in low SEP municipalities. It is important to improve the immunologic status of WLHIV as part of cancer prevention strategies in WLHIV. Cancer prevention and early detection programmes should be tailored to the needs of women ageing with HIV. In addition, SEP disparities in cancer diagnostic services have to be addressed.

KEYWORDS
epidemiology, Gynaecologic cancer, HIV, HIV RNA viral load, socioeconomic position

What's new?

Women living with HIV are at increased risk of cancers associated with human papillomavirus (HPV), such as cervical cancer. Here, the authors set out to determine the incidence and risk factors for gynaecological and breast cancers amongst women in South Africa living with HIV. Women in higher socioeconomic status municipalities were more likely to be diagnosed with breast cancer or HPV-related cancers, they found. Low CD4 counts and high HIV RNA viral loads also increased the risk of developing HPV-related cancers.

1 | INTRODUCTION

Breast and gynaecologic cancers account for approximately half of all cancers diagnosed amongst women in South Africa with breast, uterine and cervical cancer a part of the top five cancers affecting women in the country.[1] Women living with HIV (WLHIV) are at an increased risk of Human papillomavirus (HPV)-related gynaecologic cancers, such as cervical, vulvar and vaginal cancers, compared to HIV-negative women.[2,3] Although the risk of breast, ovarian and uterine cancers has not been associated with HIV, research has shown that the incidence of non-AIDS defining cancers (NADCs) has increased 3fold in recent years amongst people living with HIV.[4] This has been partially attributed to improved longevity resulting from the expanded access to antiretroviral treatment (ART) leading more WLHIV to survive long enough to develop age-related NADCs.[4,5] Besides, social determinants of health play an important role in cancer epidemiology and management, especially in low- and middle-income countries, such as South Africa.[6] Women with low socioeconomic position (SEP) have been observed to have less access to cancer care and at a higher risk of cancer compared to women with high SEP.[7]

In South Africa, HIV disproportionately affects women with a high prevalence of 17% across all ages ranging from 2.6% amongst females aged 5 to 14 years peaking at 33.3% in women aged 25 to 49 years before dropping to 13.3% in women aged 50 years and older.[8] On the other hand, breast and gynaecologic cancers make up about 50% of the new cancer cases in women. Cancer of the cervix and breast are the first and second most common causes of cancer related mortality amongst women in South Africa, respectively.[9] However, there are still gaps in information especially on the prevalence, incidence as well as risk factors for breast, ovarian and uterine cancers amongst WLHIV in South Africa. With a substantial population of women now ageing with HIV, it becomes important to understand not only HIV-related cancers but also age-related cancers in WLHIV. To assess the incidence and risk factors of cancer in people living with HIV, the South African HIV Cancer Match (SAM) study was developed.[10] This is a national HIV cohort created from routinely collected HIV data and linked to the national cancer registry data. Using this HIV cohort, we aimed to determine the incidence rate of and risk factors for developing cervical cancer and other HPV-related cancers as well as breast cancer and other hormone-related cancers in WLHIV in South Africa.

2 | METHODS

2.1 | Study design and participants

This was a longitudinal analysis of a record linkage study, including WLHIV receiving HIV care in South Africa. We used data from the SAM study, a nationwide cohort of people living with HIV.[10] The study is described in detail elsewhere.[10] Briefly, the SAM cohort is built from HIV-related laboratory records from the National Health Laboratory Service (NHLS) and cancer records from the National Cancer Registry (NCR) in South Africa. The NHLS is a network of laboratories providing its services to public sector hospitals in South Africa. The NHLS serves approximately 80% of the South African population.[11] The pathology-based NCR provided data on cancer cases. From the SAM cohort, we included women aged 15 years and older with at least two HIV-related laboratory records between 2004 and 2014.[10]

2.2 | Variables and data sources

2.2.1 | Outcome

The main outcome was breast and gynaecologic cancer diagnosis in WLHIV. Cancer cases were identified through privacy preserving probabilistic record linkages with the cancer cases recoded in the NCR. The NCR was initially established in 1986 as a pathology-based

registry, meaning it collected data for all cancers diagnosed by cytology, histology, bone marrow aspirate and trephine. In 2011, a mandate was introduced requiring both private and public healthcare facilities, including laboratories, to report diagnosed cancer cases to the NCR. The date of specimen collection was considered the date of cancer diagnosis. We estimated incidence rates of the following cancers according to the International Classification of Diseases for Oncology version 3 (ICD-O-3)[12]: cancers of the cervix (C53), other HPV-related cancers: vulva (C51), vagina (C52); breast (C50) and other hormone-related cancers: uterus (C54), ovary (C56.9); as well as other gynaecological cancers, that is, placenta (C58.9).

2.2.2 | Exposure variables

All women included in the analyses were HIV-positive. In the SAM cohort, patients were considered HIV-positive if they had test done (HIV RNA viral load or CD4 cell counts) to monitor ART or if they had a positive HIV diagnostic test (ELISA, Western Blot, or Rapid test).[10] To explore risk factors for developing HPV-related or hormone-related cancers in WLHIV, we included CD4 cell counts, HIV RNA viral loads, baseline age, calendar period at patient level; and socio-economic position (SEP) and settlement type at facility level. We used the first ever CD4 cell count or HIV RNA viral load recorded for the patient defined as the date of the first ever HIV diagnostic or monitoring test. We grouped CD4 cell counts as follows: ≤200, 201 to 350, 351 to 500 and >500 cells/μl. We grouped HIV RNA viral loads into two groups: <1000 and ≥1000 copies/ml to reflect suppressed vs unsuppressed HIV RNA viral loads as defined in the South African National HIV survey.[8] Baseline age was defined as the age at first HIV-related laboratory record in our analysis and grouped into 10-year age groups apart from those 15 to 19 years and those over 60 years. Calendar period was determined from baseline and divided to reflect the changes in HIV testing and treatment policies in South Africa: 2004 to 2007 early ART period, 2008 to 2010 middle ART period and 2011 to 2014 late ART period.[13-15] Information on individual SEP or residential address was unavailable in the NHLS laboratory records. Therefore, to determine the SEP we used information on the facility of HIV test and the associated multiple deprivation rank of the municipalities hosting the respective facilities.[16] Deprivation in this context is defined as the unmet needs of people. We used the South African Index of Multiple Deprivation, a ward level weighted aggregate of four dimensions of deprivation derived from the Statistics South Africa census data.[16] These dimensions include material, employment, education and living environment deprivation. To determine the municipal level ranking, we used population weighted average of ward ranks. The most deprived municipality (low SEP) was given a rank of one whilst the highest rank was given to the least deprived areas (high SEP). We also used the municipality information to define the health facility as rural or urban settlement. Data on municipal settlement type was determined from the national data dictionary.[17]

2.3 | Statistical methods

We described the patient characteristics stratified by prevalent, incident or no cancer. We defined prevalent cancers as those occurring on or before the first HIV-related test whilst incident cancer cases were defined as those occurring after the first HIV-related test. For patients with multiple primary cancers, we presented the first incident cancers, respectively. In addition, we described the characteristics of patients by each cancer of interest. We presented summary measures as medians and interquartile ranges for age at baseline, first CD4 cell count and person-years (pyears) of follow-up. We calculated the crude incidence rate for each cancer of interest across the 11-years study period per 100 000 pyears. We considered the date of first HIV-related laboratory records as the date of entry into the cohort (baseline). We defined the exit date from the cohort as either the date of cancer diagnosis, the date of last HIV-related test plus an additional 180 days or December 31, 2014, which was the closure of the database, whichever came first. In addition, we determined the age specific incidence rate for all cancers of interest.

To determine the factors associated with increased risk of HPV-related cancers and hormone-related cancers in WLHIV, we used Cox proportional hazard models. We specifically explored in univariable analysis the effect of CD4 cell counts, HIV RNA viral loads, baseline age, calendar period, SEP and settlement type on the incidence of cervical cancer and HPV-related cancer other than cervix (vulva and vaginal cancer), as well as breast cancer and hormone-related cancer other than breast cancer (cancer of the ovary and uterus). We excluded placenta cancers from the grouped analysis, as they did not fit the groups mentioned above. Since CD4 cell counts and HIV RNA viral loads in our study were highly correlated ($r = -0.0386$; P value <.001), we fitted separate models, including either CD4 cell counts or HIV RNA viral loads. In the multivariable analysis, the final model selection for cervical and other HPV-related as well as breast and other hormone-related cancers included age, calendar period, SEP and settlement type. CD4 cell counts and HIV RNA viral loads were only included for cervical and other HPV-related cancer. Due to the centralised cancer care model in South Africa[18] and the data linkage process,[10] we stratified all regression models by the province of first HIV-related laboratory record. We used STATA 16.1 for all analyses.

2.4 | Sensitivity analyses

Since data on HIV RNA viral load was largely missing (N = 1 309 908 [38%]), we used multiple imputation with chained equations to determine the values of the missing data assuming that the data were missing at random. We log transformed the CD4 cell counts and HIV RNA viral loads and used linear regression for the imputation of CD4 cell counts. Across the years, the detection limits of HIV RNA viral load tests have been changing. We determined that the detection limits ranged from 0 to 400 copies/ml from 2004 to 2014. However, for individuals with missing data imputed as below the detection limit of

HIV RNA viral load tests, we did not know the exact value of this missing data point. As a result, we used interval regression to impute these data. Our study included patients between the ages of 15 and 100 years, for that reason we used truncated regression to impute missing baseline age and restricted the imputed data to this age range. Therefore, any imputations beyond these limits were excluded systematically. To predict the missing data points we used patients with complete information on baseline age, HIV RNA viral loads, CD4 cell counts, calendar period, SEP and settlement type of first HIV test as well as the cancer type. We used multivariable imputation with chained equations to impute five datasets. We used the imputed dataset to determine risk factors for developing cancer and compared the results to the complete case analysis. Estimates were combined using Rubin's rules which adjusts coefficients and standard errors for variability across datasets.[19] As an additional sensitivity analysis, to assess the impact of subclinical prevalent cancers, we excluded patients who were diagnosed with cancer within the first 6 months and women that had a follow-up time of <6 months from the incidence analysis.

3 | RESULTS

3.1 | Study population

In the study period, 3 447 908 WLHIV had two or more HIV-related laboratory tests done in the public health sector. A total of 11 348 incident and 7612 prevalent hormone-related and HPV-related cancers cases were observed. Table 1 shows the characteristics of the women included in the HIV cohort. The median age at first HIV-related test was 41 years (Interquartile range [IQR]: 34-48) for WLHIV diagnosed with incident cancer and 32 years (IQR: 26-40) for WLHIV not developing cancer. Women with prevalent cancers were older at HIV diagnosis (47 years; IQR: 40-54) compared those with incident or no cancer diagnosis. The proportion of women 50 years and older with prevalent cancer was higher than that of those with incident cancers at 40% and 20%, respectively. The median first CD4 cell count was 269 cells/μl (IQR: 147-433) for WLHIV diagnosed with incident cancer and 306 cells/μl (IQR 172-474) for WLHIV not developing cancer. About 62% (n = 4809) of WLHIV diagnosed with incident cancer had a first HIV RNA viral load <1000 copies/ml, whilst 68% (n = 1 451 347) WLHIV free from cancer had an HIV RNA viral load of <1000 copies/ml. For WLHIV with and without cancer there were more patients diagnosed in high SEP municipalities compared to low SEP municipalities. We observed the same for settlement type with more WLHIV diagnosed in urban municipalities compared to rural municipalities. The median follow-up time was similar in WLHIV diagnosed with (median: 1.83 years; IQR: 0.44-3.93) and without cancer (median: 2.53 years; IQR: 1.17-4.31).

Amongst incident gynaecologic and breast cancer, cervical cancer was the most frequently diagnosis in WLHIV with 7383 cases followed by breast cancer with 2737 cases (Table 2). Placenta cancer was the least common with 52 cases. The median age at cancer diagnosis was highest for uterine cancer (51 years, IQR: 42-58) and lowest for placenta cancers (30 years, IQR: 25-36). The median CD4 cell counts were similar for all incident cancers ranging from 238 to 340 cells/μl whilst majority of WLHIV had a viral load of <1000 copies/ml across all cancers. The median follow-up time ranged from 1.60 pyears (IQR: 0.12-2.89) for placenta cancers to 2.44 pyears (0.67-4.98) for cancer of the vulva.

3.2 | Cancer incidence rates

Over 10 545 000 years of follow-up, the overall incidence rate of any gynaecological and breast cancer was 108/1 00 000 person-years (pyears). The crude incidence rate of cervical cancer ranked first with 70/1 00 000 pyears (95% confidence interval [95% CI]: 68.5-71.7), followed by breast cancer with an incidence of 26/1 00 000 pyears (95% CI: 25.0-26.9). Cancer of the placenta had the lowest incidence rate (0.49/1 00 000 pyears; 95% CI: 0.38-0.65). The age-specific incidence rate for cervical cancer rose from 3/1 00 000 pyears in 15 to 19-year-olds to 245/1 00 000 pyears in WLHIV 60 years and older (Figure 1 and Table S1). We observed a similar pattern for other HPV-related cancers. Breast cancer incidence rate ranged from 1.06/1 00 000 pyears to 120/1 00 000 pyears in the lowest and highest age groups. For other hormone-related cancers, the incidence remained low and stable in WLHIV 15 to 25 years old.

3.3 | Cancer risk factors

3.3.1 | HPV-related cancers

In the univariable and multivariable analyses, the risk of developing cervical and other HPV-related cancer decreased with increasing CD4 cell counts (Figure 2 and Tables 3 and S2). Specifically, compared to WLHIV with CD4 cell counts of <200 cells/μl, WLHIV with CD4 cell counts above 500 cells/μl had a lower risk of developing cervical cancer (Hazard ratio [HR]: 0.82; 95% CI: 0.77-0.88) and other HPV-related cancers (HR:0.62; 95% CI:0.49-0.79). Similarly, WLHIV with HIV RNA viral loads >1000 copies/ml had a higher risk of developing cervical cancer or other HPV-related cancers as compared to those with <1000 copies/ml (HR: 1.32; 95% CI: 1.25-1.40; and 1.34; 95% CI 1.11-1.63, respectively) (Figure 3 and Table S3). Older WLHIV had an increased risk of developing HPV-related cancers as compared to younger WLHIV in the models adjusting for CD4 cell counts or HIV RNA viral loads. In both models, the risk of developing cervical cancer decreased in calendar years that are more recent. In contrast, there was no such evidence for other HPV-related cancers. The risk of being diagnosed with HPV-related cancer was higher in facilities of high SEP compared to facilities in municipalities of low SEP, except for other HPV-related cancers in the model adjusting for HIV-RNA viral loads. There was no evidence of an association between HPV-related cancers risk and settlement type.

TABLE 1 Characteristics of participants by prevalent, incident and no cancer diagnosis in women living with HIV in South Africa.

	Prevalent cancers N = 7612 N (%)	Incident cancers N = 11 348 N (%)	No cancer diagnosis N = 3 436 560 N (%)
Age at first lab record (years)			
15-19	8 (0.10)	28 (0.20)	126 466 (3.9)
20-29	268 (3.50)	1237 (10.9)	1 166 213 (35.6)
30-39	1596 (21.0)	3924 (34.6)	1 142 903 (34.9)
40-49	2626 (34.5)	3639 (32.1)	561 957 (17.2)
50-59	2111 (27.7)	1928 (17.0)	219 348 (6.70)
60-69	777 (10.2)	482 (4.25)	48 426 (1.48)
70+	226 (2.97)	110 (0.97)	9389 (0.29)
Missing			161 858
Median (IQR)	47 (40-54)	41 (34-48)	32 (26-40)
First CD4 cell count recorded (cells/μl)			
≤200	2411 (32.3)	4097 (36.5)	1 021 535 (30.2)
201-350	2009 (26.9)	3094 (27.5)	930 480 (27.5)
351-500	1339 (17.9)	1980 (17.6)	684 006 (20.2)
≥501	1705 (22.8)	2065 (18.4)	748 082 (22.1)
Missing	148	112	52 457
Median (IQR)	298 (163-479)	269 (147-433)	306 (172-474)
First HIV RNA viral load (copies/ml)			
<1000	3214 (70.3)	4809 (61.6)	1 451 347 (68.0)
≥1000	1361 (29.7)	3000 (38.4)	681 881 (32.0)
Missing	3037	3539	1 303 332
Calendar period of first laboratory record			
2004-2006	841 (11.0)	3158 (27.8)	491 273 (14.3)
2007-2010	3254 (42.7)	5491 (48.4)	1 509 644 (43.9)
2011-2014	3517 (46.2)	2699 (23.8)	1 435 643 (41.8)
Facility related municipality characteristics			
Socioeconomic position			
Low	815 (10.7)	944 (8.30)	545 357 (15.9)
Lower-middle	1270 (16.7)	1720 (15.2)	572 936 (16.7)
Upper-middle	1333 (17.5)	2062 (18.2)	583 361 (17.0)
High	4191 (55.1)	6612 (58.3)	1 732 420 (50.4)
Missing	3	10	2486
Settlement type			
Rural	3791 (49.8)	5308 (46.8)	1 855 369 (54)
Urban	3818 (50.2)	6030 (53.2)	1 578 705 (46)
Missing	3	10	2486

3.3.2 | Hormone-related cancers

The strongest predictor for developing hormone-related cancers in WLHIV was age. Older WLHIV had an increased risk of developing hormone-related cancer compared to younger WLHIV, with the effect more pronounced than for HPV-related cancers. There was also some evidence of an increased risk of breast cancer in facilities of high municipal SEP compared to facilities of low municipal SEP. Settlement type was not associated with hormone-related cancers risk. The results for the individual cancers are presented in the supplement (Tables S10-S12).

3.3.3 | Findings from sensitivity analyses

With the multiple imputation analyses, we observed results similar to the complete case analyses across all cancer groups (Tables S7-S9). Likewise, when we restricted the follow-up time to those with

TABLE 2 Characteristics of women living with HIV at first HIV-related test stratified by each female specific cancer.

	N (%)						
	Cervix N = 7383	Breast N = 2737	Vagina N = 143	Vulva N = 465	Uterus N = 302	Ovary N = 173	Placenta N = 52
Patient related characteristics							
Age at first lab record (years)							
15-19	12 (0.20)	4 (0.10)		2 (0.40)	2 (0.7)	3 (1.7)	5 (9.6)
20-29	797 (10.8)	234 (8.50)	19 (13.3)	124 (26.7)	11 (3.6)	31 (17.9)	18 (34.6)
30-39	2676 (36.2)	894 (32.7)	52 (36.4)	189 (40.6)	46 (15.2)	34 (19.7)	19 (36.5)
40-49	2416 (32.7)	925 (33.8)	46 (32.2)	94 (20.2)	79 (26.2)	49 (28.3)	10 (19.2)
50-59	1161 (15.7)	519 (19.0)	21 (14.7)	49 (10.5)	109 (36.1)	43 (24.9)	0
60-69	260 (3.50)	141 (5.20)	5 (3.50)	5 (1.10)	42 (13.9)	13 (7.50)	0
70+	61 (0.80)	20 (0.70)	0	2 (0.40)	13 (4.30)	0	0
Median (IQR)	40 (34-47)	42 (35-49)	40 (32-48)	34 (29-43)	51 (42-58)	43 (33-52)	30 (25-36)
First CD4 cell count recorded (cells/μl)							
≤200	2782 (38.1)	870 (32.1)	57 (39.9)	191 (41.4)	98 (32.8)	62 (36)	14 (27.5)
201-350	1991 (27.2)	765 (28.2)	39 (27.3)	129 (28.0)	86 (28.8)	52 (30.2)	13 (25.5)
351-500	1270 (17.4)	519 (19.1)	28 (19.6)	63 (13.7)	42 (14.0)	27 (15.7)	14 (27.5)
≥501	1266 (17.3)	559 (20.6)	19 (13.3)	78 (16.9)	73 (24.4)	31 (18)	10 (19.6)
Missing	74	24		4	3	1	1
Median (IQR)	263 (141-422)	292 (163.4-465)	248 (141-426.25)	238 (138-386)	280 (165-492)	273 (145-413)	340 (189-468)
First HIV RNA viral load (copies/ml)							
<1000	3029 (60.5)	1247 (63.8)	66 (64.7)	206 (58.9)	134 (65.4)	75 (63.0)	13 (52.0)
≥1000	1977 (39.5)	709 (36.2)	36 (35.3)	144 (41.1)	71 (34.6)	44 (37.0)	12 (48.0)
Missing	2377	781	41	115	97	54	27
Calendar period of first lab record							
2004-2006	2053 (27.8)	809 (29.6)	38 (26.6)	142 (30.5)	64 (21.2)	51 (29.5)	10 (19.2)
2007-2010	3535 (47.9)	1346 (49.2)	61 (42.7)	216 (46.5)	160 (53.0)	87 (50.3)	29 (55.8)
2011-2014	1795 (24.3)	582 (21.3)	44 (30.8)	107 (23)	78 (25.8)	35 (20.2)	13 (25)
Facility related municipality characteristics							
Socioeconomic position							
Low	605 (8.2)	224 (8.20)	14 (9.8)	40 (8.60)	39 (12.9)	10 (5.80)	3 (5.90)
Lower-middle	1186 (16.1)	365 (13.4)	20 (14)	54 (11.6)	49 (16.2)	25 (14.5)	8 (15.7)
Upper-middle	1403 (19.0)	459 (16.8)	20 (14)	80 (17.2)	49 (16.2)	29 (16.8)	10 (19.6)
High	4186 (56.7)	1684 (61.6)	89 (62.2)	290 (62.5)	165 (54.6)	109 (63.0)	30 (58.8)

(Continues)

TABLE 2 (Continued)

	N (%)						
	Cervix N = 7383	Breast N = 2737	Vagina N = 143	Vulva N = 465	Uterus N = 302	Ovary N = 173	Placenta N = 52
Missing	3	5		1		1	11
Settlement type							
Rural	3582 (48.5)	1179 (43.2)	64 (44.8)	199 (42.9)	145 (48.0)	73 (42.2)	24 (47.1)
Urban	3798 (51.5)	1553 (56.8)	79 (55.2)	265 (57.1)	157 (52.0)	100 (57.8)	27 (52.9)
Missing	3	5		1		1	
Median follow-up time (pyears)	1.61 (0.32-3.72)	2.34 (0.86-4.23)	2.17 (0.33-3.85)	2.44 (0.67-4.98)	1.71 (0.28-3.84)	1.63 (0.38-3.70)	1.60 (0.12-2.89)
Crude incidence rate/1 00 000 pyears (95%CI)	70.0 (68.5-71.7)	25.9 (25.0-26.9)	1.35 (1.15-1.59)	4.40 (4.02-4.82)	2.86 (2.55-3.20)	1.64 (1.41-1.90)	0.49 (0.38-0.65)

Abbreviations: 95% CI, 95% confidence interval; IQR, interquartile range.

>6 months follow-up time, we observed results comparable to the complete case analysis (Tables S4 and S5).

4 | DISCUSSION

In this South African nationwide study of gynaecological and breast cancers in WLHIV, we observed the highest incidence rate for cervical cancer followed by cancer of the breast, vulva, uterus, ovary, vagina and placenta. We observed a higher risk of breast, cervical and other HPV-related cancer diagnosis in WLHIV accessing HIV care in high SEP areas compared to women accessing care in areas of low SEP. The risk of developing cervical cancer or another HPV-related cancer increased with decreasing CD4 cell counts and increasing HIV RNA viral loads. Cancer risk in WLHIV increased with older age for all cancer types studied with more pronounced effects for hormone-related cancers.

Our study is the first to explore breast and all gynaecologic cancers and their risk factors in WLHIV at a population level in South Africa. Since the HIV cohort was created using NHLS data, our study had a population coverage of about 80%. It is also the first study to look at the association of municipal SEP of health facilities and risk of breast and all gynaecologic cancers other than cervical cancer in WLHIV in South Africa. Our study had some limitations. Our HIV data might be incomplete with point of care HIV tests largely missing in the NHLS database. We observed a relatively short follow-up time in our cohort. In addition, the NCR estimates in our study were from the pathology-based registry, leading to under ascertainment of cancers that are diagnosed only clinically. Therefore, our incidence rates may under- or overestimate the true burden of cancer. The under- or overestimates may also differ by cancer type, as some cancers are diagnosed more easily than others. A large proportion of participants had missing HIV RNA viral loads, which we accounted for using multiple imputation. Still, we had limited information to predict the missing data, which might result in limitations similar to the complete case analysis. Data on ethnicity, which has been shown to be associated with disparities in cancer risk amongst people living with HIV, was unavailable. However, the assumption would be that most of the women were black as the prevalence of HIV is highest amongst black women in South Africa.[8] We lacked information on other risk factors than age for hormone-related cancers than age such as, obesity and hormonal contraceptive use, which might have resulted in residual confounding thus the large effect of age observed. Data on SEP and settlement type was at municipal level and not at individual level. Therefore, interpretation of findings should be kept at municipal level, as we cannot infer individual risk.

The incidence rates that we observed in WLHIV were lower than that observed in WLHIV in South Africa and Zimbabwe[20-22] but also higher than reported in Malawi.[23] The inconsistencies in the reported incidence rates could have been due to differences in study designs and definitions of time at risk. As explained in the limitations section, we cannot exclude that shortcomings in the estimation of our study numerator and denominator may have led to an underestimation of

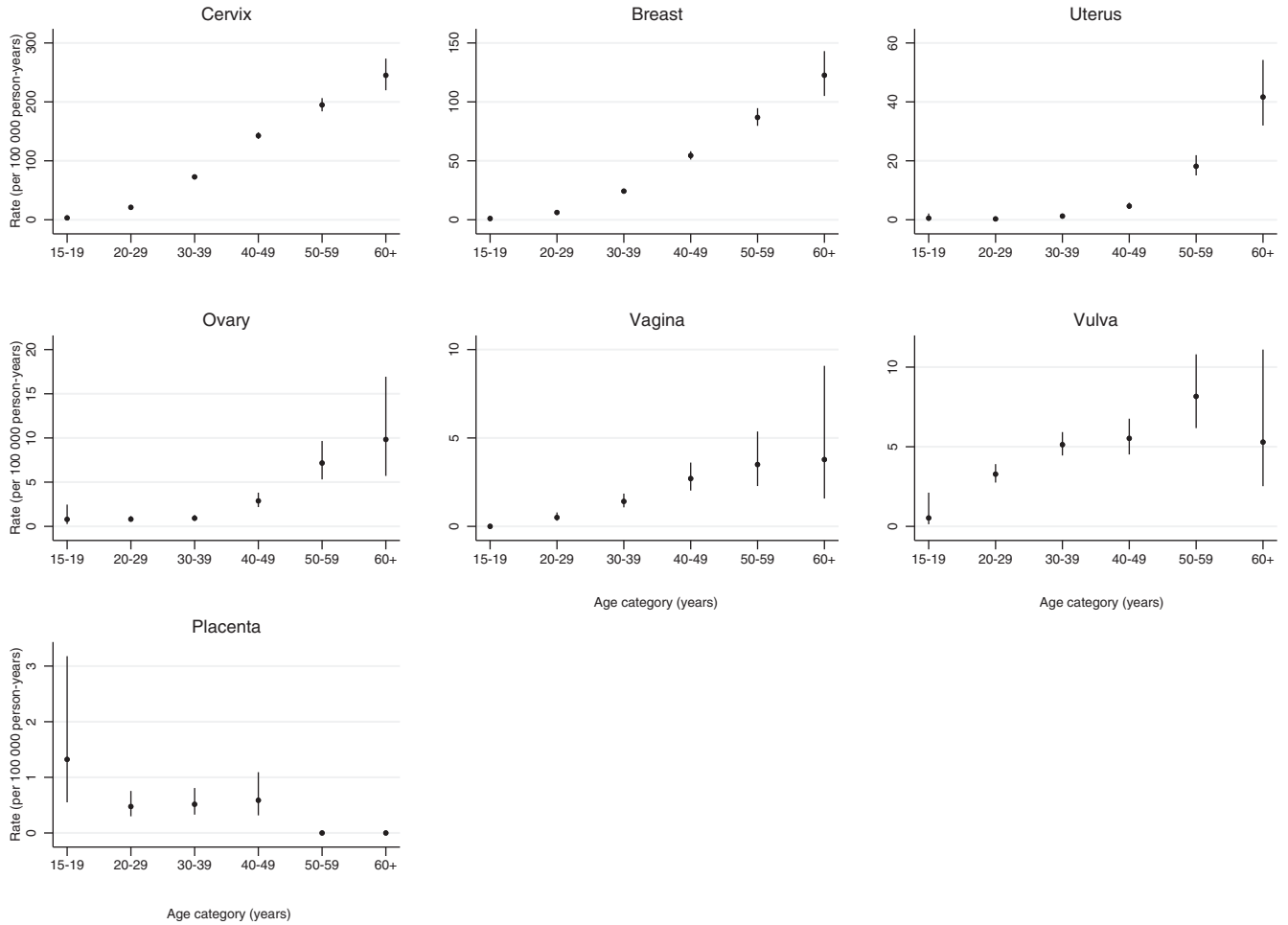

FIGURE 1 Age specific incidence rate/100000 person-years by individual cancer.

the reported cancer incidence rates. In our study, we observed an elevated risk of HPV-related cancers with decreasing CD4 cell counts and high-level HIV RNA viremia. Our results are in line with retrospective studies documenting low median CD4 cell counts at the time of cancer diagnosis for HPV-related cancers including vulvar and cervical cancer, and higher CD4 cell counts in other hormone-related cancers, that is, ovary, uterine cancer.[4] In a multicentre cohort study on WLHIV diagnosed with gynaecologic cancers, high-level viremia was only observed for cervical cancer.[4] The increased incidence of HPV-related cancers in WLHIV has been attributed to HIV-induced immunodeficiency, elevated HPV prevalence and persistence of HPV-infections.[24-27]

Age is an established risk factor for cancer development, and in our study, older ages were associated with a higher risk of all cancers under study with the strongest association observed for other hormone-related cancers. A study including WLHIV in the United States observed that the median age for vulvar, cervical, ovary and uterine cancer was 47, 45, 50 and 53 years, respectively[4] and this is older than the 34, 40 and 43 and 51 years we observed for the same cancers in our cohort of WLHIV. A younger HIV population in South Africa compared to other regions might explain these differences.[28] Another analysis form the SAM study evaluated the association of age and cancer in people living with HIV in South Africa and showed that whilst infection related cancers were more common, infection unrelated cancers were, predominant in HIV patients aged 54 years and older compared to younger people.[29] Few studies have assessed the incidence and age-specific incidence rate of other hormone-related cancers in WLHIV globally and in African settings these data are generally not available for comparison. Of note is the number of prevalent cancers in WLHIV aged 50 and older which even exceeded the incident cancers reported in the same age groups. This corresponds to the national HIV survey results which show that, about 54% of women aged 50 years and older are unaware of their HIV positive status.[8]

The diagnosis of breast, cervical and other HPV-related cancer in our cohort was associated with health facilities in high SEP municipalities. Using area level deprivation, most studies in the general European population have demonstrated an elevated risk of breast cancer and a lower risk of cervical cancer in areas of high SEP compared to areas of low SEP.[30-33] In a previous analysis of the SAM study evaluating the spatiotemporal distribution of cervical cancer in WLHIV across municipalities, an elevated incidence of cervical cancer in high SEP municipalities as well as municipalities with a higher number of health facilities was observed.[34] In our study, the increased risk

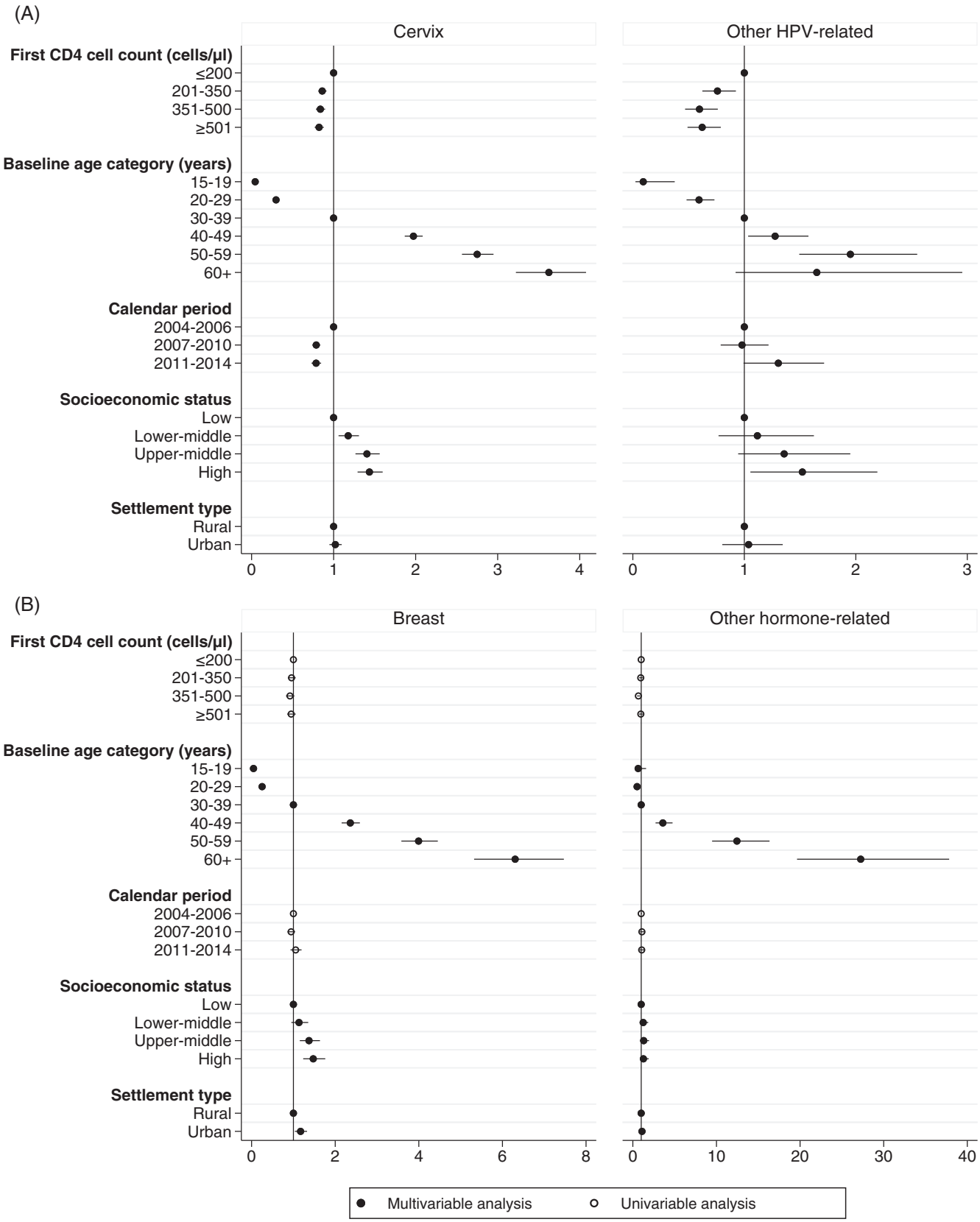

FIGURE 2 Factors associated with gynaecologic and breast cancer in WLHIV: models including CD4 cell count. Cervical cancer and other HPV-related cancer models adjusted for baseline age, calendar period, socio-economic status and settlement type (A). Breast and other hormone-related cancer models adjusting for baseline age, socio-economic status and settlement type (Panel B). Other HPV related = cancer of the vagina and vulva. Other hormone related = cancer of the uterus and ovary. Clear circle represents variables only evaluated in the univariable analysis.

TABLE 3 Univariable analysis of the factors associated with cancer incidence in women living with HIV.

	HPV-related cancer		Hormone-related cancer	
	Cervix HR (95% CI)	Other HR (95% CI)	Breast HR (95% CI)	Other HR (95% CI)
Patient level characteristics				
First CD4 cell count recorded (cells/μl)				
≤200	1	1	1	1
201-350	0.79 (0.74-0.83)	0.73 (0.60-0.88)	0.96 (0.87-1.05)	0.94 (0.75-1.19)
351-500	0.71 (0.66-0.75)	0.55 (0.44-0.71)	0.92 (0.82-1.02)	0.66 (0.50-0.88)
≥501	0.67 (0.63-0.72)	0.56 (0.44-0.71)	0.95 (0.85-1.05)	0.95 (0.74-1.22)
First HIV RNA viral load (copies/ml)				
<1000	1	1	1	1
≥1000	1.25 (1.18-1.32)	1.28 (1.05-1.55)	1.05 (0.96-1.16)	1.04 (0.82-1.31)
Age at first lab record (years)				
15-19	0.05 (0.03-0.08)	0.08 (0.02-0.34)	0.05 (0.02-0.12)	0.64 (0.26-1.58)
20-29	0.29 (0.27-0.32)	0.57 (0.47-0.71)	0.25 (0.22-0.29)	0.51 (0.35-0.74)
30-39	1	1	1	1
40-49	1.98 (1.87-2.09)	1.30 (1.06-1.60)	2.36 (2.15-2.59)	3.58 (2.71-4.74)
50-59	2.71 (2.53-2.91)	1.95 (1.49-2.54)	3.94 (3.54-4.40)	12.4 (9.42-16.3)
60+	3.49 (3.11-3.92)	1.65 (0.92-2.96)	6.14 (5.18-7.27)	26.9 (19.4-37.4)
Calendar period of first lab record				
2004-2006	1	1	1	1
2007-2010	0.86 (0.81-0.91)	1.01 (0.82-1.26)	0.95 (0.86-1.04)	1.09 (0.85-1.39)
2011-2014	0.92 (0.86-0.99)	1.35 (1.03-1.77)	1.05 (0.93-1.19)	1.06 (0.78-1.44)
Facility related municipality characteristics				
Socioeconomic position				
Low	1	1	1	1
Lower-middle	1.16 (1.05-1.29)	1.11 (0.76-1.61)	1.12 (0.94-1.34)	1.18 (0.80-1.74)
Upper-middle	1.41 (1.28-1.56)	1.37 (0.96-1.95)	1.40 (1.19-1.66)	1.23 (0.84-1.80)
High	1.44 (1.30-1.58)	1.58 (1.15-2.17)	1.53 (1.31-1.79)	1.17 (0.82-1.66)
Settlement type				
Rural	1	1	1	1
Urban	1.15 (1.08-1.23)	1.28 (1.03-1.58)	1.33 (1.19-1.47)	1.10 (0.86-1.41)

Abbreviations: CI, confidence interval; HR, hazard ratio; HPV related, HPV-related cancer other than cervix, specifically cancer of the vagina and vulva; Hormone related, cancer of the uterus and ovary.

of HPV-related and breast cancer diagnosis in high compared to low SEP municipalities most likely reflects the centralization of South African cancer care.[35] This subsequently leads to higher laboratory confirmed cancer diagnoses. In addition the high proportion of specialised health practitioners particularly in facilities in high SEP and urban municipalities in the country might result in higher cancer detection rates in municipalities of high SEP.[36] Another study evaluating the effect of life course SEP on breast and cervical cancer screening rates indicated higher screening rates in participants with higher SEP.[37] The increased screening rates and higher access to cancer care in high SEP municipalities can also result in better cancer detection rates in these areas. Similar to other studies in the general population we observed no evidence of an association between municipality level SEP with ovarian and uterine cancers.[32] We did not detect a significant association between rural and urban settlement types and cancer risk. A case control study in India showed an increased risk in cervical cancer in rural areas compared to urban areas.[38] In addition, it has been noted that women from rural areas have less access to information on cancer and health care related services, factors that are associated with increased cancer risk. A study in the American general population also observed that women in rural areas likely had less access to gynaecology oncologists resulting in elevated incidence of gynaecologic cancer.[39]

Cancer surveillance in WLHIV is becoming increasingly important especially in the context of women ageing with HIV as it allows for better understanding of cancer burden in this population including risk factors, prognostic factors as well as treatment options. More studies investigating the epidemiology, prevention, diagnosis and treatment

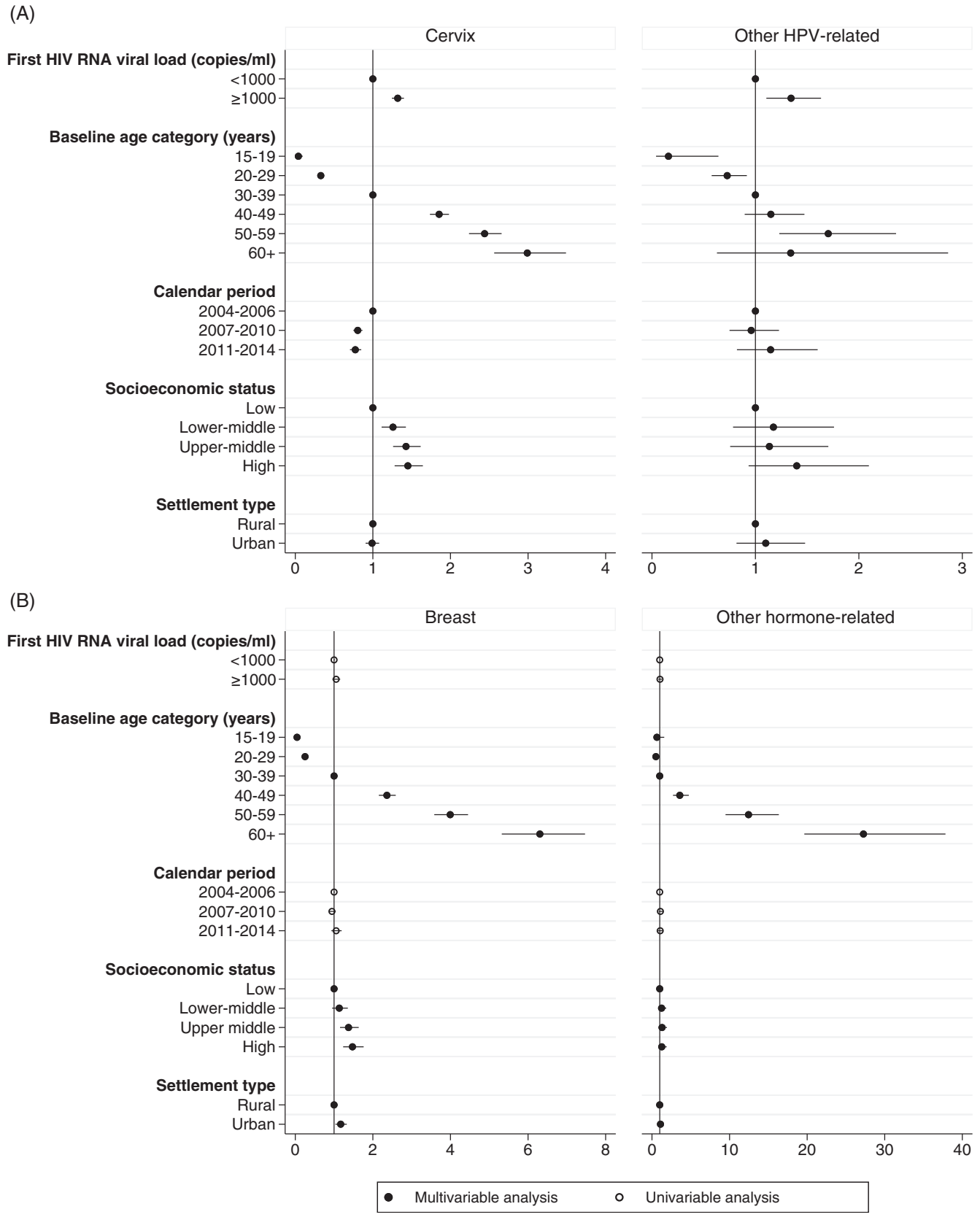

FIGURE 3 Factors associated with gynaecologic and breast cancer in WLHIV: models including HIV RNA viral loads. Cervical cancer and other HPV-related cancer models adjusted for HIV RNA viral load, baseline age, calendar period, socio-economic status and settlement type (A). Breast and other hormone-related cancer models adjusting for baseline age, socio-economic status and settlement type (B). Other HPV related = cancer of the vagina and vulva. Other hormone related cancer = cancer of the uterus and ovary. Clear circles represent variables only evaluated in the univariable analysis. A, HPV-related cancers; B, hormone-related cancers.

of gynaecologic cancer, particularly hormone-related cancers, should be encouraged. Women should continue to be encouraged to go for HIV testing across all age groups. Integration of HIV and other women's health services should still be encouraged and go beyond cervical cancer screening. Strategies such as cancer screening, early detection (where applicable), and diagnostic services should be scaled-up to reach women at high risk and in low SEP areas. Given that cervical and other HPV-related cancers contribute a significant proportion of the cancer burden in WLHIV, effective interventions like HPV-vaccination and cervical screening using HPV tests remain essential components in the prevention and early detection of these cancers.

5 | CONCLUSIONS

In conclusion, low CD4 cell counts and high HIV RNA viral loads increased the risk of developing HPV-related cancers amongst WLHIV. Older age was a risk factor for all cancers under study in WLHIV. HPV-related cancers and breast cancer diagnosis was associated with facilities in high SEP municipalities. Whilst cancer prevention and early detection programmes should consider women ageing with HIV, it remains important to improve the immunologic status of WLHIV and address SEP disparities in cancer burden.

AUTHOR CONTRIBUTIONS

Tafadzwa G. Dhokotera, Maša Davidović, Matthias Egger, Mazvita Muchengeti and Julia Bohlius contributed towards the study design. Tafadzwa G. Dhokotera contributed towards literature search, data analysis and drafting of first version of manuscript. Mazvita Muchengeti contributed towards data acquisition. Victor Olago contributed towards data linkage. All authors contributed towards data interpretation and critical comments on the first and subsequent drafts of the manuscript. The work reported in the paper has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

The authors would like to thank the late Dr. Elvira Singh for her contributions towards the South African Match study. They thank the National Health Laboratory Service and the National Cancer Registry for making the data available. They would also like to than Dr. Jan Hattendorf for the statistical input. Open access funding provided by Universitat Basel.

FUNDING INFORMATION

This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (NIH) and by the National Cancer Institute (NCI) under award number U01AI069924 to the IeDEA-Southern Africa (ongoing sub-award), an NIH Administrative Supplement (grant number U01AI069924-09 to Matthias Egger and Julia Bohlius), PEPFAR supplement (to Matthias Egger), the Swiss National Science Foundation (SNSF, grant number 320030_169967 to Julia Bohlius and Matthias Egger), and the US CRDF Global (grant number HIV_DAA3-16-62705-1 to Mazvita Muchengeti). Matthias Egger was supported by special project funding (grant number 207285) from the SNSF, and Tafadzwa G. Dhokotera and Maša Davidović by the European Union's Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement number 801076), through the SSPH+ Global PhD Fellowship Programme.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The original data set analyzed in the present article used patient-level data and was provided by the Academic Affairs and Research Office at the National Health Laboratory Service (NHLS), South Africa. Data is available from the NHLS (Contact academic.research@nhls.as.za) for researchers who meet the relevant criteria for access to these data. Further information is available from the corresponding author upon request.

ETHICS STATEMENT

This study was approved by the Human Research Ethics Committee of the University of the Witwatersrand (M190594), Johannesburg, South Africa, and the Cantonal Ethics committee (2016-00589) in Bern, Switzerland.

ORCID

Tafadzwa G. Dhokotera https://orcid.org/0000-0002-4353-0787
Mazvita Muchengeti https://orcid.org/0000-0002-1955-923X
Eliane Rohner https://orcid.org/0000-0002-0554-2875

TWITTER

Tafadzwa G. Dhokotera *GladysDhokotera*

REFERENCES

1. National Cancer Registry. Cancer in South Africa 2019. 2021.
2. Stein L, Urban MI, O'Connell D, et al. The spectrum of human immunodeficiency virus-associated cancers in a South African black population: results from a case-control study, 1995-2004. *Int J Cancer*. 2008;122:2260-2265.
3. Dryden-Peterson S, Medhin H, Kebabonye-Pusoentsi M, et al. Cancer incidence following expansion of HIV treatment in Botswana. *PLoS One*. 2015;10:1-13.
4. Levinson KL, Riedel DJ, Ojalvo LS, et al. Gynecologic cancer in HIV infected women: treatment and outcomes in a multi-institutional cohort. *Aids*. 2018;32:171-177.
5. Cubasch H, Joffe M, Hanisch R, et al. Breast cancer characteristics and HIV among 1,092 women in Soweto, South Africa. *Breast Cancer Res Treat*. 2013;140:177-186.
6. Gelband H, Sankaranarayanan R, Gauvreau CL, et al. Costs, affordability, and feasibility of an essential package of cancer control interventions in low-income and middle-income countries: key messages from disease control priorities, 3rd edition. *Lancet*. 2016;387:2133-2144.
7. Ginsburg O, Bray F, Coleman MP, et al. The global burden of women's cancers: an unmet grand challenge in global health. *Lancet*. 2017;389:847-860.
8. Simbayi LC, Zuma K, Zungu N, et al. South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2017. 2019.

9. Ferlay J, Ervik M, Lam F, et al. *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer; 2020.
10. Muchengeti M, Bartels L, Olago V, et al. Cohort profile: the South African HIV cancer match (SAM) study, a national population-based cohort. *BMJ Open*. 2022;12:e053460.
11. Blecher M, Kollipara A, DeJager P, Zulu N. Health financing. In: Padarath A, English R, eds. *South African Health Review*. Health Systems Trust. Durban, South Africa; 2011.
12. Fritz A, Percy C, Jack A, et al. *International Isabel Classification of Diseases for Oncology*. Geneva, Switzerland. 3rd ed.; 2013.
13. Southern African HIV Clinicians Society. Guidelines for adherence to antiretroviral therapy in adolescents and young adults (expanded version): Recommendations, resources and references. 2017.
14. South Africa National Department of Health. South Africa antiretroviral treatment guidelines. 2004.
15. Meyer-Rath G, Johnson LF, Pillay Y, et al. Changing the South African national antiretroviral therapy guidelines: the role of cost modelling. *PLoS One*. 2017;12:e0186557.
16. Noble M, Zembe W, Wright G, Avenell D. *Multiple Deprivation and Income Poverty at Small Area Level in South Africa in 2011*. Cape town, South Africa; 2013.
17. National Department of Health. National Department of Health Data Dictionary. 2019.
18. National Department of Health. Cervical Cancer Prevention and Control Policy. 2017.
19. Rubin D. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, Ltd. New York, USA; 1987.
20. Sengayi M, Spoerri A, Egger M, et al. Record linkage to correct under-ascertainment of cancers in HIV cohorts: the Sinikithemba HIV clinic linkage project. *Int J Cancer*. 2016;139:1209-1216.
21. Rohner E, Sengayi M, Goeieman B, et al. Cervical cancer risk and impact of Pap-based screening in HIV-positive women on antiretroviral therapy in Johannesburg, South Africa. *Int J Cancer*. 2017;141:488-496.
22. Shamu T, Rohner E, Chokunonga E, et al. Cancer incidence among people living with HIV in Zimbabwe: a record linkage study. *Cancer Rep*. 2022;5:e1597.
23. Horner M-J, Chasimpha S, Spoerri A, 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.
24. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *The Lancet*. 2007;370:890-907.
25. Stelzle D, Tanaka LF, Lee KK, et al. Estimates of the global burden of cervical cancer associated with HIV. *Lancet Glob Health*. 2021;9:e161-e169.
26. Wielgos AA, Pietrzak B. Human papilloma virus-related premalignant and malignant lesions of the cervix and anogenital tract in immunocompromised women. *Ginekol Pol*. 2020;91:32-37.
27. Ruffieux Y, Muchengeti M, Egger M, et al. Immunodeficiency and cancer in 3.5 million people living with human immunodeficiency virus (HIV): the South African HIV cancer match study. *Clin Infect Dis*. 2021;73:e735-e744.
28. McCormack VA, Febvey-Combes O, Ginsburg O, dos Santos Silva I. Breast cancer in women living with HIV: a first global estimate. *Int J Cancer*. 2018;143:2732-2740.
29. Ruffieux Y, Muchengeti M, Olago V, et al. Age and cancer incidence in 5.2 million people with human immunodeficiency virus (HIV): the South African HIV cancer match study. *Clin Infect Dis*. 2022;76:1440-1448. doi:10.1093/CID/CIAC925
30. Jensen KE, Hannibal CG, Nielsen A, et al. Social inequality and incidence of and survival from cancer of the female genital organs in a population-based study in Denmark, 1994-2003. *Eur J Cancer*. 2008;44:2003-2017.
31. Arik A, Dodd E, Cairns A, Streftaris G. Socioeconomic disparities in cancer incidence and mortality in England and the impact of age-at-diagnosis on cancer mortality. *PLoS One*. 2021;16:e0253854.
32. Mihor A, Tomsic S, Zagar T, Lokar K, Zadnik V. Socioeconomic inequalities in cancer incidence in Europe: a comprehensive review of population-based epidemiological studies. *Radiol Oncol*. 2020;54:1-13.
33. Hoebel J, Kroll LE, Fiebig J, et al. Socioeconomic inequalities in total and site-specific cancer incidence in Germany: a population-based registry study. *Front Oncol*. 2018;8:402.
34. Tafadzwa D, Julien R, Lina B, et al. Spatiotemporal modelling and mapping of cervical cancer incidence among HIV positive women in South Africa: a nationwide study. *Int J Health Geogr*. 2021;20:30.
35. South African National Department of Health. National Cancer Strategic Framework for South Africa 2017-2022. 2017.
36. Kawonga M, Fonn S. Achieving effective cervical screening coverage in South Africa through human resources and health systems development. *Reprod Health Matters*. 2008;16:32-40.
37. Akinyemiju T, Ogunsina K, Sakhuja S, Ogbhodo V, Braithwaite D. Life-course socioeconomic status and breast and cervical cancer screening: analysis of the WHO's study on global ageing and adult health (SAGE). *BMJ Open*. 2016;6:e012753.
38. Kashyap N, Krishnan N, Kaur S, Ghai S. Risk factors of cervical cancer: a case-control study. *Asia Pac J Oncol Nurs*. 2019;6:308-314.
39. Ackroyd SA, Shih YCT, Kim B, Lee NK, Halpern MT. A look at the gynecologic oncologist workforce: are we meeting patient demand? *Gynecol Oncol*. 2021;163:229-236.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dhokotera TG, Muchengeti M, Davidović M, et al. Gynaecologic and breast cancers in women living with HIV in South Africa: A record linkage study. *Int J Cancer*. 2024;154(2):284-296. doi:10.1002/ijc.34712

B-cell malignancies - A new knowledge hub on the latest research in therapeutic advances

EDUCATIONAL CONTENT AVAILABLE ON THE HUB:

- On-demand Webinars - earn CME credit
- Infographics
- Patient Case Studies
- Currated Research Articles

...and much more

VISIT KNOWLEDGE HUB TODAY

This educational resource has been supported by Eli Lilly.

WILEY