The Oncologist, 2026, **31(1)**, oyaf408
https://doi.org/10.1093/oncolo/oyaf408
Advance access publication 13 December 2025
Original Article

Overall survival, poverty differentials, and mediating pathways among women with breast cancer: South African Breast Cancer and HIV Outcomes cohort

Maureen Joffe, PhD[1,2,*], Wenlong Carl Chen, PhD[1,2,3,4], Ashleigh Craig, PhD[5], Daniel S. O'Neil, MD[6], Alfred I. Neugut, MD[7,8,9], Judith S. Jacobson[8], Paul Ruff, MD[1,10,11], Rofhiwa Margaret Mathiba, MD[11], Nivashini Murugan, MD[12], Herbert Cubasch, MD[12], Charmaine Blanchard, MD[1], Sarah Nietz, PhD[12], Jennifer Edge, MD[12], Ines Buccimazza, MD[13], Sharon Čačala, MD[14], Valerie A. McCormack, MD[15], Yoanna Pumpalova, PhD[7,9], Shane A. Norris, PhD[5,16]

[1]Strengthening Oncology Services Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, 2193, South Africa
[2]Network for Oncology Research in Africa (NORA), Global Health Working Group, Martin-Luther-University, Halle-Wittenberg, 06888, Germany
[3]National Cancer Registry, a Division of the National Institute for Communicable Diseases, National Health Laboratory Service, Johannesburg, 2000, South Africa
[4]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, 2193, South Africa
[5]SAMRC/Wits Developmental Pathways for Health Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, 2193, South Africa
[6]Yale Cancer Center, Yale School of Medicine, Yale University, Connecticut, CT, 06511, United States
[7]Herbert Irving Comprehensive Cancer Center, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, 10032, United States
[8]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, 10032, United States
[9]Department of Medicine, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, United States
[10]South Africa Medical Research Council Common Epithelial Cancers Research Centre, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[11]Division of Medical Oncology, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, 2193, South Africa
[12]Department of Surgery, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, 2193, South Africa
[13]Department of Specialized Surgery, Inkosi Albert Luthuli Central Hospital, Durban and Ngwelezane Hospital, University of KwaZulu-Natal, Empangeni, 3880, South Africa
[14]Division of Surgery, Stellenbosch University and Tygerberg Hospital, Cape Town, 7500, South Africa
[15]Environment and Lifestyle Epidemiology Branch, International Agency for Research on Cancer, Lyon, 69366, France
[16]School of Human Development and Health, University of Southampton, Southampton, SO17 1BJ, United Kingdom

*Corresponding author: Strengthening Oncology Services Research Unit, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg, 2193, South Africa (mjoffe@witshealth.co.za).

Abstract

Background: Breast cancer survival rates in sub-Saharan Africa are low. In a prospective, multi-center cohort study, we estimated 5-year overall survival rates, overall survival determinants, and mediating effects between socioeconomic status on overall survival among South African women diagnosed with invasive BC.

Patients and methods: Patients from 4 public hospitals were enrolled between July 1, 2015 and January 31, 2019. Survival determinants were assessed using Cox proportional hazard models adjusted for age, background mortality, and treatments. Socioeconomic pathway effects on overall survival were determined through generalized structural equation models.

Results: Of 2838 participants, 58% had advanced-stage (III/IV) disease. Five-year crude overall survival was 44.3% (95% CI 42.5-46.2). Significant mortality risks were late stage at diagnosis (hazard ratio [HR] = 2.31 [95% CI 1.99-2.69] [stage III]; 4.79 [95% CI 3.96-5.80] [stage IV]), HIV-positive status (HR = 1.45 [95% CI 1.25-1.67]), unemployment HR = 1.25 [95% CI 1.09-1.44], and low education HR 1.19 [95% CI 1.04-1.37]). Age and treatment-adjusted socioeconomic status effects on overall survival were mediated through HIV status (81.7% of the effect) and stage at

Received: 12 August 2025; Accepted: 7 December 2025.

diagnosis (81.7%), both $P < .001$. Poor breast cancer knowledge had an indirect effect on overall survival, accounting for 77.6% of the total effect ($P = .001$), fully mediated by late-stage presentation. Socioeconomic status had no significant direct path to mortality after accounting for these mediators.

Conclusion: Interventions should prioritize early breast cancer detection. For patients with low socioeconomic status, particularly those with comorbid HIV, we must mitigate multifaceted barriers to healthcare access, including limited awareness and knowledge of breast cancer.

Key words: breast cancer; overall survival; determinants and socioeconomic mediators; South Africa.

Implications for Practice

In the South African Breast Cancer and HIV Outcomes study cohort, 5-year crude overall survival was a low 44%, despite free access to cancer diagnostics and treatments for patients treated within the Public Health Sector. Major survival determinants were late stage at diagnosis, HIV-positive status, and socioeconomic vulnerability. The generalized structural equation model revealed that socioeconomic status effects on overall survival were mediated through HIV status, breast cancer knowledge, and timely diagnosis among socioeconomically vulnerable women, after adjusting for age and treatment effects. Interventions must address both structural and informational barriers to early diagnosis and care.

Introduction

Breast cancer is the most common cancer in women (2.3 million diagnosed in 2022; ~12% of all cancers in both sexes) and the leading cause of cancer death among them (670,000 deaths in 2022).[1] High-human development index (HDI) countries bear the highest incidence (age-standardized rate (ASIR) with ranges of 80 to >100 per 100,000 population), around double that for low- and middle-HDI countries (28 to >55 ASIR).[1,2] However, the incidence is rapidly increasing in low- and middle-HDI countries, which contribute over 53% of all new global breast cancer cases.[3] Given the relatively lower number of incident cancers, deaths from cancer and mortality to incidence ratios are disproportionately high in low- and middle-HDI countries, which contribute around 62% of global breast cancer mortality.[4] By 2050, new cases and deaths are expected to increase by 38% and 68%, respectively, disproportionately affecting countries with low HDI.

The sub-Saharan African region has among the highest breast cancer mortality to incidence ratios globally, with 5-year overall survival a low 40% compared with 85%-90% in HIC settings.[5,6] If not addressed, the region, along with other low- and middle-HDI settings, will bear the brunt of the global breast cancer burden in years to come. A recent systematic review of the sub-Saharan African regions' breast cancer studies[6] revealed respective 1, 2, 3, 4, and 5-year survival rates with 95% CIs of 79% (67%-88%), 70% (57%-80%), 56% (45%-67%), 54% (43%-65%), and 40% (32%-49%). Encouragingly, the same review showed improved survival patterns over time and significant variation in reported incidence, morbidity, and mortality among populations from sub-Saharan Africa, reflecting varying levels of patient access to screening and early diagnostic services and cancer treatments.[6,7] Furthermore, poverty is a critical upstream determinant of health that intersects with numerous downstream factors influencing breast cancer outcomes. In low- and middle-income settings, as well as among socioeconomically disadvantaged populations in high-income countries, poverty shapes the entire cancer care continuum, from early detection to treatment completion and follow-up care.[8] Poverty is associated with lower health literacy, reduced screening uptake, and delayed help-seeking, often leading to diagnosis at more advanced stages, a key predictor of poorer survival.[9,10] It is thus necessary to understand modifiable upstream sociodemographic and downstream health system factors that negatively impact breast cancer survival to inform national, regional, and international policy interventions.[11]

In 2015, we established the South African Breast Cancer and HIV Outcomes (SABCHO) longitudinal cohort study to prospectively collect robust phenotypic and long-term vital status data from South African women newly diagnosed with invasive breast cancer and cared for in the Public Health Sector. From this cohort, we have previously identified that having greater socioeconomic status, fewer children, greater knowledge, and awareness of breast cancer, along with a family history of breast cancer and waiting less than 3 months to patient first health system access, were the upstream sociodemographic factors positively associated with early-stage breast cancer diagnosis. In contrast, having aggressive breast cancer subtypes and having to negotiate convoluted journeys to diagnostic centers involving secondary hospitals were determinants of late stage at diagnosis, which is a critical intermediate outcome that affects survival.[9,12] It is well known that social determinants of health are at the root of South Africa's HIV epidemic and poor health outcomes.[13] Yet contrary to findings in other settings,[14] in our cohort, which has a 22% prevalence of HIV, we found that HIV status had no significant impact on stage at breast cancer diagnosis.[12] Therefore, the aim of this paper is 3-fold: first, to assess 5-year overall survival levels, second, to identify the up- and downstream factors associated with overall survival; and third, to determine direct and indirect effects that poverty may have on 5-year overall survival among breast cancer patients from socioeconomically disadvantaged communities in South Africa. Additionally, by modeling poverty not only as a covariate but also through mediation analysis, we can uncover the mechanisms through which poverty exerts its effect, whether via knowledge, late-stage diagnosis, or HIV infection. This deepens our understanding of how structural factors translate into clinical outcomes, helping to identify intervention points that improve survival equitably.

Methods

Setting

South Africa is a middle-HDI country, where high levels of inequality, unemployment, and poverty persist, adversely affecting some 80% of the population.[15] The country has dual healthcare systems, the wealthiest 15% of the population is privately insured; the remaining 85% are dependent on the resource-constrained public health system.[16] The South Africa public health system provides cancer diagnostic and treatment services at almost no cost to patients, except for out-of-pocket transport and hospital visit costs.[16] This is unlike in most countries in sub-Saharan Africa, where patients must fund their own diagnostic and treatment costs.

Study design

Between July 1, 2015 and January 31, 2019, SABCHO prospectively enrolled 2838 economically disadvantaged South Africa women 18 years of age or older, with stage I-IV invasive breast cancer, newly diagnosed and treated at 4 Academic Hospitals located in Gauteng and KwaZulu-Natal Provinces, the most populated provinces of the country, which also bear high HIV prevalence burdens, into a prospective longitudinal cohort. The protocol, study design and data variable collection[17] was guided by the Engel classic bio-psychosocial theoretical model.[18,19] The sample size for the SABCHO cohort was calculated to achieve 80% power, a type 1 error rate of 5%, and a 2-sided log-rank test enabling detection of differences in survival as small as 10%, given the HIV prevalence of the cohort at around 22%. The 4 tertiary hospital sites selected were the Chris Hani Baragwanath Academic Hospital, in Soweto, located in southern Johannesburg, named the "Soweto" site, the Charlotte Maxeke Johannesburg Academic Hospital, in central Johannesburg, named the "Johannesburg" site, the Inkosi Albert Luthuli Central Hospital and Addington Hospital, located in Durban, KwaZulu-Natal collectively called the "Durban" site, and the Greys Hospital site, in Pietermaritzburg, KwaZulu-Natal called the "Pietermaritzburg" site. The sites serve a range of urban, peri-urban, and rural communities, which are representative of the communities within their respective catchment areas. Participant and site details have been previously described.[19]

Outcome data

The primary outcome was overall survival assessed at routine follow-up visits. This endpoint was selected because it is clinically meaningful and could be most reliably measured in our resource-constrained settings. Patients were assessed at 3-6 monthly routine clinic follow-up visits, and those who missed their scheduled clinic visits were contacted every 3 months to determine their vital status.[20] If the patient, next of kin, and other persons named as close contacts could not be reached for 2 consecutive follow-up calls, we searched VerifyID (www.verifyid.co.za), a publicly available administrative database, to determine the patient's vital status. Patient survival was censored at the last date they were known to be alive if no additional information about vital status could be obtained (3%). Among those known to have died, date of death information was 74.2% from next of kin, 5.3% from hospital records, and 20.5% from VerifyID.

Statistical analysis

We calculated overall survival starting from the histopathologically confirmed date of diagnosis by core biopsy. Follow-up continued until either the date of death, the date on which the participant was last known to be alive, 5 years after diagnosis, or December 31, 2023, whichever came first. Crude Kaplan–Meier survival curves were used to visually represent the results. We calculated net overall survival, accounting for background age-specific mortality for South Africa, and estimated age-standardized net overall survival to the International Cancer Survival Standard.[7] The 5-year age-specific death rates in 2019 for South African women were sourced from the World Health Organization Global Health Observatory Life tables and expanded to death rates by single year of age using the flexible Poisson model of the WHO.[21] Overall survival determinants were assessed using Cox proportional hazard models stratified by individual and cumulative domains. The combined model adjusted for age at diagnosis, background mortality, and treatment because our primary interest was in factors present before or at diagnosis and their impact on overall survival.

In preliminary analyses (Table S3, Model 7), we observed a survival difference by Hospital Site. However, further assessment showed that the effect of Site was almost entirely accounted for by treatments received. Consequently, we excluded Site from subsequent analyses and included treatments received in the multivariable Cox regression models as a covariate rather than a factor of primary interest. This approach allowed us to adjust for the potential confounding influence of treatment received while assessing the associations between other variables of interest and overall survival. Subsequent testing showed strong collinearity between Hospital Site and Clinical Stage at Diagnosis. We evaluated this relationship using likelihood ratio tests between nested Cox models and pairwise correlation analysis (pwcorr). The likelihood ratio test indicated no significant improvement in model fit when including the interaction term, and the correlation between Hospital Site and Stage suggested overlapping explanatory effects. We also used paramed and medeff packages in Stata to explore whether Stage at Diagnosis mediated the association between Hospital Site and survival. Both parametric and g structural equation model-based mediation models (paramed, gsem) confirmed that Stage largely explained the effect of Hospital Site on survival. Given the high collinearity and mediation through Stage, Hospital Site was excluded from the final models to avoid overadjustment and unstable estimates. While a hierarchical (multi-level) model could theoretically account for site-level variance, the small number of hospitals ($n = 4$) provided insufficient cluster-level degrees of freedom to justify a random-effects specification. Sensitivity analyses using fixed effects for site produced comparable estimates, supporting the robustness of the final model specification. This approach allowed us to adjust for the potential confounding influence of treatment received while assessing the associations between other variables of interest and overall survival.

A generalized structural equation model was used to assess direct and indirect effects of socioeconomic wealth status on the mortality of breast cancer patients (adjusted for age, background mortality, and treatments received) and further explored mediation effects of HIV status, breast cancer knowledge, and stage at diagnosis. The wealth status was based on

a composite score of 6 household possessions (Home ownership = 1, Car ownership = 1, Microwave = 1, Washing machine = 1, Indoor hot & cold running water = 1, Flush toilet inside home = 1). A score of 0-3 was categorized as low to middle wealth status. The direct (unmediated), indirect (mediated), and total effects of the model were computed and recorded, and the proportion of the total effect mediated was calculated. Modifications to pathways and adding/removing variables were made iteratively, and the Akaike and Bayesian Information Criteria (IC)[22] of each model were compared. The final model was selected for having the lowest IC and high theoretical relevance. Direct, indirect, and total effects were calculated using non-linear combination estimates.

All analyses were performed using Stata SE version 17 (StataCorp). Cox proportional hazards regression models were fitted using the stcox command to estimate hazard ratios (HRs) with 95% CIs. Model assumptions were tested using Schoenfeld residuals (estat phtest) to verify proportional hazards. Influential observations were examined using deviance residuals. Kaplan–Meier survival curves were generated with the sts graph command to visualize survival probabilities and compare groups using log-rank tests.

Generalized structural equation modeling was conducted using Stata's gsem command to assess hypothesized pathway relationships among variables. For binary outcome variables, the logit link function was applied to model the probability of the event occurring, and coefficients were expressed as odds ratios (ORs). For continuous outcomes, the identity link was used. Model fit was evaluated using likelihood ratio tests, Akaike's Information Criterion, and Bayesian Information Criterion. Standardized coefficients were reported to facilitate comparison across pathways.

Data missingness

The data collection was robust, with low missingness (Table S1). Missing data were addressed using multiple imputation by chained equations in Stata SE version 17. The dataset was set to mlong format, and relevant variables were registered for imputation. We generated 20 imputations with a burn-in of 10 iterations to ensure convergence and stable parameter estimates. Predictive mean matching (pmm, knn(5)) was applied for continuous variables, binary variables were imputed using the logit model, and categorical variables with more than 2 levels were imputed using the multinomial logit (mlogit) model. All analyses following imputation were performed using the mi estimate prefix, which automatically combines results across imputations using Rubin's rules to account for both within- and between-imputation variance. Parameter estimates are presented as exponentiated coefficients (hazard ratios, HRs) with 95% CIs.

We considered best- and worst-case assumptions to investigate whether the missing data patterns occurred randomly. Our findings (Figure S1, see online supplementary material for a color version of this figure) revealed no material changes in the Cox regression model hazard ratios for the complete-case analysis, with, as expected, slightly tighter confidence interval ranges. This suggests randomness of missing data from these 2 variables and robustness of the data as a whole.

Role of the funding source

Funding sources had no role in study design, data collection, data analysis, data interpretation, or writing of the article.

Results

Between July 1, 2015 and January 31, 2019, of 2974 women examined for eligibility, we excluded 110 who had in situ disease, 9 women with breast sarcomas, and 17 with phyllodes tumors. We included 2838 women newly diagnosed with stage I-IV invasive breast carcinomas (Figure S2, see online supplementary material for a color version of this figure). Descriptive characteristics and survival data at each site and overall for the study population are summarized in Table 1 and provided in more detail in Tables S1 and S2, The mean age at diagnosis was 55.9 years, with participants from Soweto and Johannesburg sites presenting relatively younger compared to those from Durban, and Pietermaritzburg. The Pietermaritzburg site had the highest proportion of low to middle wealth scores (73.4%), minimal social support (69.6%), unemployment (80.2%), and poor to intermediate knowledge and awareness of breast cancer (63.9%). Close to 14% of women had family members diagnosed with breast cancer. Some 21.9% of enrolled women were HIV-positive at diagnosis and the majority of participants (58.3%) were diagnosed with advanced disease (stage III) (40.1%) and stage IV (18.3%) disease. Immunohistochemistry determined subtype distribution was 60.2% for hormone receptor (HR)-positive, human epidermal growth factor receptor 2 (HER2)-negative tumors, 17.0% for HR-positive, HER2-positive tumors, 6.6% for HR-negative, HER2-positive, and 15.6% were triple negative. Overall, 7.7% of participants received no treatment, 2.4% received surgery with no systemic treatments, 10.1% received neoadjuvant/palliative chemotherapy, 7.9% received endocrine treatment only (reflecting old age and frailty), and the rest received treatment modalities based on staging and receptor subtype considerations.

Of 2838 women followed for 5 years, 1555 (55%) died, 1283 (45%) were alive at the 5-year administratively censored time period, 33 (1.2%) were censored early (before the 5-year follow-up period), and 59 participants (2.1%) were lost to follow-up (Table 1 and Table S2). Median age at death was 56 years (IQR 45-68). Overall, 15.1% died within 1 year of diagnosis and crude 5-year overall survival was 44.3%, with minor variations between hospital sites. Age-standardized net 5-year survival estimates at 51.6% did not differ substantially from the net survival values (Table 1 and Table S2).

As shown in crude Kaplan–Meier curves in Figure 1A and B (with age, background mortality, and treatment-adjusted hazard ratios and 95% CIs), breast cancer stage at diagnosis was a strong prognostic factor, with 5-year crude survival at 65% in patients with stage I+II disease vs 33.9% in those with stage III and 7.7% with stage IV disease. Women with comorbid HIV infection had lower 5-year crude survival (45% for HIV-negative versus 33% for HIV-positive participants). As shown in the forest plot (Figure 2), in the combined model adjusted for age, background mortality, and treatments received, besides stage at diagnosis and HIV status other factors significantly associated with poorer overall survival were, being single, unemployed, having informal or primary school education only, having one or more self-reported cardiovascular disease (CVD) comorbidities, and tumors with a high Ki-67 proliferation index. In contrast, having a family member diagnosed with breast cancer was associated with increased survival.

In the domain-specific analyses (Table S3), the association between treatment and overall survival (Model 7) persisted, even after adjustment for stage at diagnosis and receptor subtype in the fully adjusted model. In contrast, the apparent site

Table 1. Characteristics, deaths, and survival estimates of the SABCHO cohort by enrolment site.

Exposure domain *N* (%)	Total	Soweto	Johannesburg	Durban	Pietermaritzburg
Number of women enrolled and followed by *N* (%)	2838 (100)	1024 (36)	685 (24)	603 (21)	526 (19)
Age at diagnosis: Mean (*SD*)	55.9 (14.3)	55.0 (14.6)	54.8 (14.0)	57.0 (14.0)	57.6 (14.4)
Household and sociodemographic factors					
Single (including divorced and widowed) vs cohabiting	1754 (61.8)	607 (59.3)	394 (57.5)	387 (64.2)	366 (69.6)
Unemployed (including retired and students)	2062 (72.7)	746 (72.9)	437 (63.8)	457 (75.8)	422 (80.2)
Primary school or less education (R0/G7 or informal) vs secondary or higher	873 (30.8)	274 (26.8)	144 (21.0)	229 (38.0)	226 (43.0)
Low to Mid wealth index score (0-3) vs high score[a]	1556 (54.8)	514 (50.2)	302 (44.1)	354 (58.7)	386 (73.4)
Family member diagnosed with cancer	394 (13.9)	118 (11.5)	103 (15.0)	117 (19.4)	56 (10.6)
Poor to intermediate knowledge of breast cancer signs and symptoms (score 0-5)[b]	1565 (55.1)	566 (55.3)	364 (53.1)	299 (49.6)	336 (63.9)
Clinical factors					
HIV status at time of breast cancer diagnosis					
Positive	621 (21.9)	251 (24.5)	120 (17.5)	112 (18.6)	138 (26.2)
Stage at diagnosis					
Stages I+II	1177 (41.8)	511 (49.9)	256 (37.2)	207 (34.3)	204 (38.8)
Stage III	1139 (40.1)	386 (37.7)	318 (46.4)	243 (40.1)	192 (36.5)
Stage IV	519 (18.3)	126 (12.3)	112 (16.4)	151 (25.0)	130 (24.7)
Cancer proliferation risk					
High: Ki67≥20%	2006 (70.7)	804 (78.5)	461 (67.3)	406 (67.3)	335 (63.7)
Receptor subtype risk					
HR+/HER2-	1709 (60.2)	586 (57.4)	411 (60.0)	377 (62.5)	333 (63.3)
HR+/HER2+	483 (17.0)	223 (21.8)	105 (15.3)	91 (15.1)	64 (12.2)
HR-/HER2+	186 (6.6)	53 (5.2)	55 (8.0)	45 (7.5)	33 (6.3)
HR-/HER2-	444 (15.6)	155 (15.1)	110 (16.1)	87 (14.4)	92 (17.5)
Death and survival data					
Median time since diagnosis IQR, years[c]	3.8 (1.6-5.0)	4 (1.7-5.0)	3.5 (1.5-5.0)	4.1 (1.8-5.0)	3.8 (1.4-5.0)
Died at end of follow-up[d]	1555 (54.8)	541 (52.8)	378 (55.2)	328 (54.4)	308 (58.6)
1-year crude survival[e]	84.9 (83.5-86.1)	86.3 (84.0-88.2)	82.1 (79.8-85.5)	87.7 (84.8-90.1)	82.5 (79.0-85.5)
3-year crude survival[e]	58.2 (56.3-60.0)	59.0 (56.1-62.1)	57.4 (53.6-61.1)	59.6 (55.6-63.4)	56.1 (51.7-60.2)
5-year crude survival[e]	44.3 (42.8-46.5)	46.0 (42.9-49.0)	43.2 (39.4-46.9)	45.4 (41.4-49.4)	41.3 (37.1-45.5)

Treatment details, other variables and missing data (generally well below 10%) are provided in Table S1. Missing data were minimal, generally well below 10%. Full details are provided in Table S1.
Abbreviations: IQR = interquartile range; SD = standard deviation.
[a]Wealth score determined from: (Home ownership = 1, Car ownership = 1, Microwave = 1, Washing machine = 1, Indoor hot & cold running water = 1, Flush toilet inside home = 1. Denominator = 6 (1 for Yes, 0 for No).
[b]Self-reported knowledge and awareness of breast cancer. 1 was allocated for every correct answer, 0 for wrong answer or don't know, as previously described.[23] HR (Estrogen receptor/Progesterone receptor), HER2 (Human Epidermal Growth Factor Receptor 2).
[c]Regardless of vital status.
[d]End of follow-up to earliest of 5 years after diagnosis or December 31, 2023, whichever came first.
[e]Data are *n* (%) or percentage surviving (95% CI) unless otherwise indicated.

effect was completely attenuated once treatment was taken into account, indicating that differences in the treatment received explain inter-site differences in survival.

Lessons learned are that crude 5-year overall survival was very low, associated with late stage at diagnosis, HIV-positive status, low knowledge of breast cancer, and poverty. Apparent site survival differences were explained by differing treatments received.

The generalized structural equation model (Figure 3, Tables 2 and 3) revealed that socioeconomic status (household possession wealth) had a significant total effect on 5-year mortality following breast cancer diagnosis, largely mediated by HIV status and stage at diagnosis. The direct effect of wealth on mortality was not statistically significant, indicating that the association between lower socioeconomic status and higher mortality operates primarily through indirect pathways. HIV status emerged as a strong mediator, with a significant indirect effect, accounting for 81.7% of the total effect of socioeconomic status on mortality. Similarly, stage at diagnosis mediated 81.7% of the effect of socioeconomic status on mortality, suggesting that women from lower socioeconomic status backgrounds were more likely to be diagnosed at a later stage, significantly increasing mortality risk. In contrast, poor breast cancer knowledge had a smaller and non-significant mediating role, contributing to a modest total effect without evidence of significant mediation. Further analysis showed that poor breast cancer knowledge increased mortality through its strong indirect effect on late-stage diagnosis with a total effect of 0.424 (95% CI: 0.167-0.681, $P = .001$), indicating that 77.6% of the effect was mediated by stage. We examined multicollinearity between HIV and late stage, which were significantly associated and added a path link in the model between HIV and stage to account for this. This improved the model fit but did not substantially alter the mediation results.

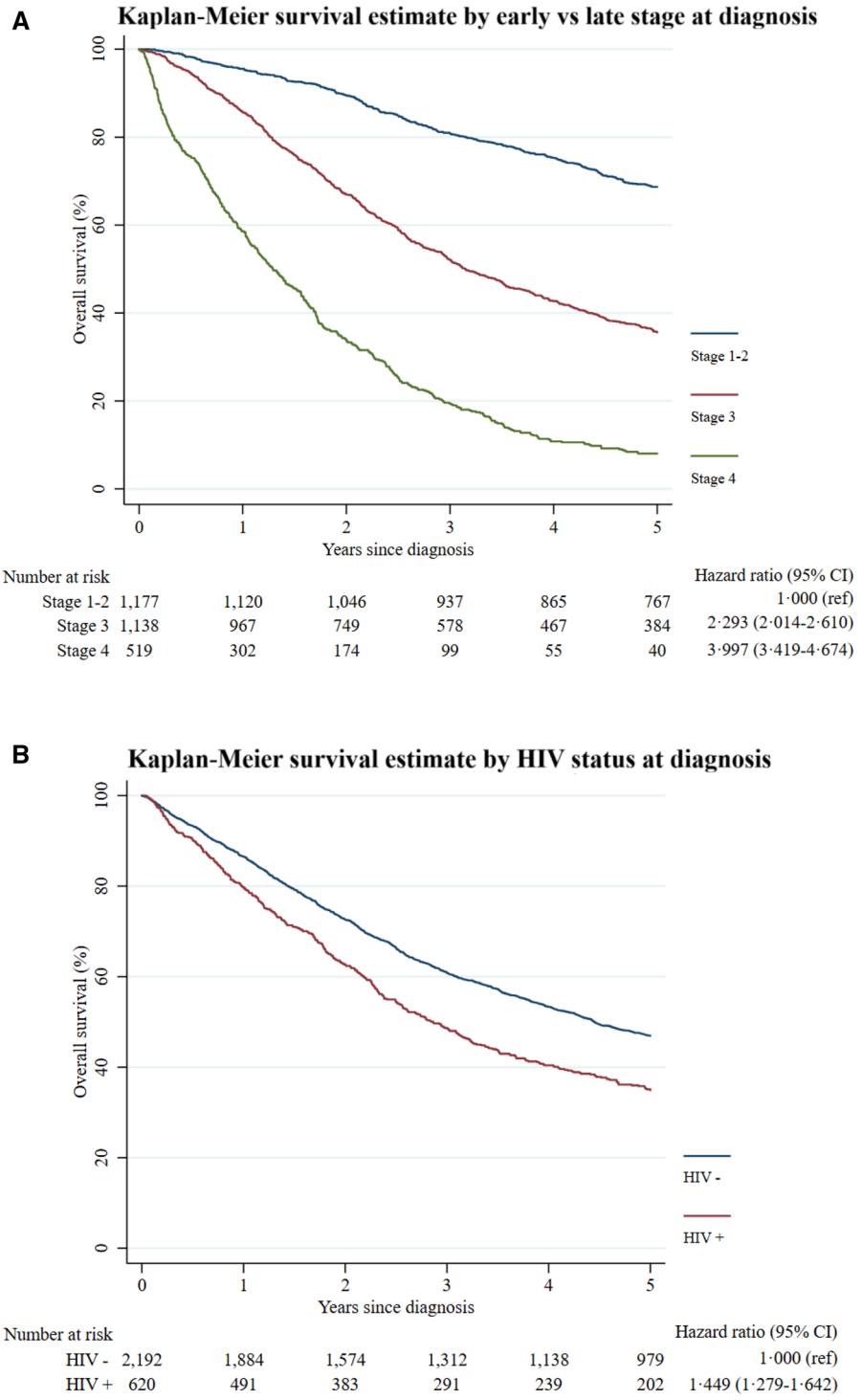

Figure 1. Kaplan–Meier survival curves (unadjusted) showing crude 5-year overall survival in 2838 women by (A) stage at diagnosis and (B) by HIV status. Hazard ratios adjusted for age, background mortality rates, and treatment received.

In parallel mediation models (Table 2), each mediator's indirect effect is calculated independently of the others, assuming no mediation path overlaps. However, the results indicate that the mediators may not operate via distinct, non-overlapping mechanisms and that there is possibility of shared variance. We consequently executed a joint mediation model (Table 3) and found that HIV status and late stage remained mediators of the relationship between wealth and mortality, accounting for 42.5% each of the total effect, while breast cancer knowledge showed no significant mediation.

Lessons learned were that after adjustment for age and treatment effects, socioeconomic differences in mortality are largely explained by disparities in people living with HIV and late stage at diagnosis.

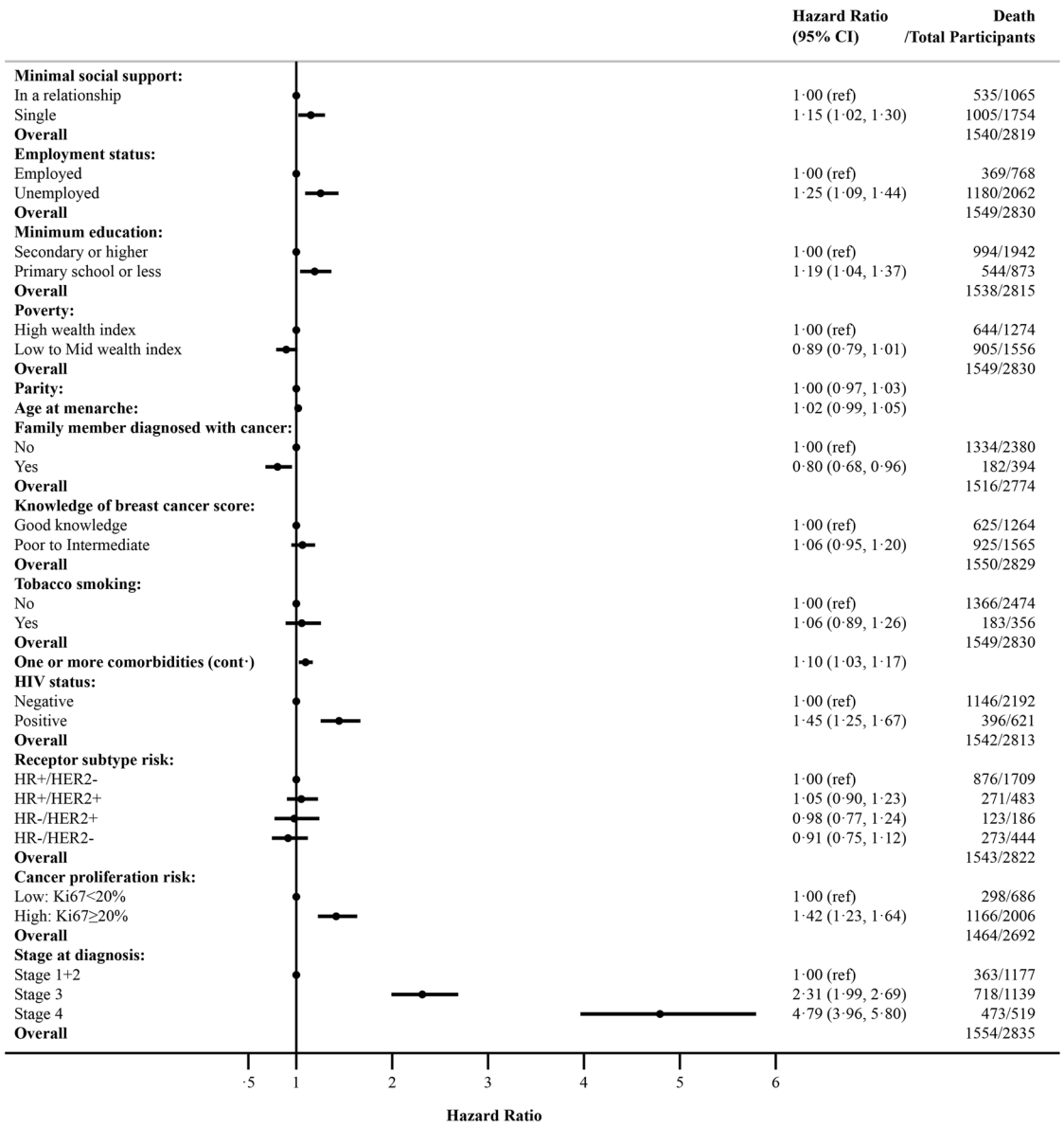

Figure 2. Age and background mortality adjusted HRs for 5-year all-cause mortality among 2838 women, in the combined adjusted model by socioeconomic characteristics, behavioral and reproductive risks, family history of breast cancer, breast cancer knowledge, HIV and cardiovascular disease risk, and cancer clinical characteristics. HR+ = estrogen receptor-positive or progesterone receptor-positive; HR- = estrogen receptor-negative and progesterone receptor-negative; HER2 = human epidermal growth factor receptor *2*. Hazard ratios from Cox model adjusted for age at diagnosis, background mortality rate, and treatments received.

Discussion

This study provides the first robust, population-based evidence on 5-year overall survival for women diagnosed with breast cancer within South Africa's public health sector. We observed that 15% of patients died within the first year and that crude 5-year survival was 44% increasing to 52% when age-standardized. These survival rates are markedly lower than those reported in high-HDI countries and within South Africa's private sector, where population-wide mammography screening and access to timely, high-quality treatments contribute to 5-year survival rates of 85%-90%.[23] Our findings parallel outcomes among socioeconomically marginalized populations in HICs[24–26] and other sub-Saharan African contexts.[7,27–29] As modeled by McCormack et al., achieving earlier diagnosis and universal treatment access in these settings could yield survival gains of up to 20%.[7]

Stage at diagnosis effects on overall survival

Within the SABCHO cohort, stage at diagnosis emerged as the most significant determinant of mortality—even after adjusting for age, background mortality, and treatment received. Women diagnosed with stage III disease had more than double the risk of death (HR 2.31), while stage IV disease conferred nearly a 5-fold increase (HR 4.79).

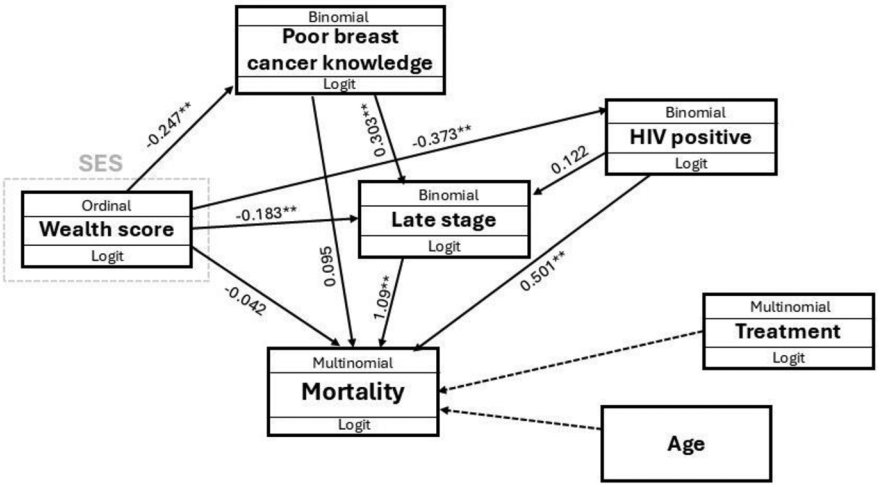

Figure 3. Generalized structural equation model for SES, knowledge and HIV status. SES = socioeconomic status. *$P < .05$; **$P \leq .00$.

Table 2. Adjusted generalized structural equation model (parallel mediation model) in a sample of respondents for socioeconomic position, knowledge, and HIV status.

			Direct effect (95% CI)	*P* value	Indirect effect (95% CI)	*P* value	Total effect (95% CI)	*P* value	Proportion mediated
Effect of SES on mortality via HIV status									
Wealth	Mortality	HIV (positive)	−0.042 (−0.097; 0.012)	.13	−0.187 (−0.280; −0.094)	**<.001**	−0.229 (−0.328; −0.131)	**<.001**	81.7%*
Effect of SES on mortality via breast cancer knowledge									
Wealth	Mortality	Knowledge (poor)	−0.042 (−0.097; 0.012)	.13	−0.024 (−0.069; 0.022)	.31	−0.066 (−0.129; −0.002)	**.043**	–
Effect of SES on mortality via stage									
SES	Mortality	Stage (late)	−0.042 (−0.097; 0.012)	.13	−0.187 (−0.252; −0.122)	**<.001**	−0.229 (−0.312; −0.147)	**<.001**	81.7%*
Effect of breast cancer knowledge on mortality via stage									
Knowledge (poor)	Mortality	Stage (late)	0.095 (−0.087; 0.277)	.31	0.332 (0.147; 0.517)	**<.001**	0.428 (0.170; 0.685)	**.001**	77.6%*

Models were adjusted for age and treatment. Abbreviations: SES – socioeconomic status. *Direct, indirect and total effects significant at p< 0.05. The proportions mediated are full mediation effects

Table 3. Adjusted generalized structural equation model in a sample of respondents for socioeconomic position, knowledge, and HIV status showing a total effect accounting for all 3 mediators.

			Direct effect (95% CI)	*P* value	Indirect effect (95% CI)	*P* value	Total effect (95% CI)	*P* value	Proportion mediated
Effect of SES on mortality via HIV status									
Wealth	Mortality	HIV (positive)	−0.042 (−0.097; 0.012)	.13	−0.187 (−0.280; −0.094)	**<.001**	−0.440 (−0.561; −0.319)	**<.001**	42.5%[a]
Effect of SES on mortality via breast cancer knowledge									
Wealth	Mortality	Knowledge (poor)	−0.042 (−0.097; 0.012)	.13	−0.024 (−0.069; 0.022)	.31	−0.440 (−0.561; −0.319)	**<.001**	–
Effect of SES on mortality via stage									
SES	Mortality	Stage (late)	−0.042 (−0.097; 0.012)	.13	−0.187 (−0.252; −0.122)	**<.001**	−0.440 (−0.561; −0.319)	**<.001**	42.5%[a]

Models were adjusted for age and treatment.
Abbreviation: SES = socioeconomic status.
[a]Direct, indirect and total effects significant at $P < .05$. Proportions mediated are full mediation effects.

HIV and co-morbidity effects on overall survival

HIV co-infection was independently associated with a 45% increased risk of death despite widespread access to antiretroviral therapy. The presence of one or more comorbidities associated with CVD risk were associated with a 10% increased risk in death.

Household and socioeconomic effects on overall survival

Additional predictors of poor survival included low socioeconomic status, lack of family, or household exposure to breast cancer (suggesting limited awareness) and the presence of comorbidities associated with CVD risk. Collectively, these findings reflect the profound influence of social determinants of health on cancer outcomes and echo global evidence linking structural disadvantage to disparities in cancer survival.[30,31]

Lessons learned are that when population-based mammography screening is not feasible, survival rates remain extremely low. In SABCHO, the major survival determinant is late stage at diagnosis but also poor knowledge and awareness of early breast cancer symptoms, poverty, HIV-positive status, cardiovascular risk comorbidity burden, and socioeconomic disadvantage.

Socioeconomic mediation effects on overall survival

Our generalized structural equation modeling further elucidated the pathways through which socioeconomic status influences overall survival. Poverty adversely impacted survival via 3 main pathways: advanced stage at diagnosis, HIV-positive status, and reduced breast cancer knowledge; the latter exerting its effect entirely through late-stage disease. We did observe collinearity between HIV status and late-stage presentation and accounted for this in structural equation models, as it reflects the intertwined biological and social mechanisms through which HIV infection may influence tumor progression. Lessons learned are that these findings should be interpreted as reflecting partially overlapping pathways, where HIV contributes both directly to poorer outcomes and indirectly through later stage at diagnosis, but the collinearity did not attenuate the mediation effect of late-stage presentation. The results further suggest shared real-world overlap and variance across the mediating pathways, particularly the higher likelihood of late-stage diagnosis among women living with HIV.

These findings underscore that in resource-constrained environments like South Africa, where diagnostic and treatment services are ostensibly available, structural inequities and knowledge gaps remain critical barriers to survival. Interventions should consider extending beyond facility-based care to community-level strategies that raise awareness of early signs and symptoms of breast cancer, facilitate prompt care-seeking, and support navigation through the health system. Importantly, women with HIV and breast cancer represent a clinically vulnerable subgroup who warrant tailored care approaches. Integrated oncologic and HIV care models are urgently needed, including HIV specialists in weekly oncology multidisciplinary treatment planning meetings, to consider drug interactions, immunologic vulnerabilities, and tailored treatment monitoring. National policy guidelines should be accordingly revised to address the specific needs of this high-risk group, particularly as the population ages and breast cancer incidence continues to rise among women living with HIV.

Implications for practice

Meeting the WHO Global Breast Cancer Initiative targets—diagnosing 60% of invasive breast cancer at stages I-II, including a pathologically confirmed diagnosis within 60 days of first health facility access, and ensuring 80% of patients complete quality treatments—could improve breast cancer survival in South Africa.[7,32] In the absence of affordable population-based mammography screening, community-based education, symptom awareness, and navigation strategies are essential to promote earlier stage at diagnosis. Concurrently, policy makers and health providers must prioritize women with HIV and breast cancer for integrated, person-centered care to optimize survival outcomes. Our findings should inform both national cancer control planning and broader health system strengthening efforts that aim to address inequities in cancer outcomes.

Limitations and strengths

This study relies on self-reported socioeconomic status and comorbidity data, which may introduce recall or reporting bias. Moreover, while the study sites represent the 2 most populous provinces, generalizability to other regions in South Africa may be limited. Nonetheless, the strengths of this work lie in its prospective cohort design, comprehensive sociodemographic and clinical data, low rates of loss to follow-up (2%), and minimal data missingness, making it the most comprehensive dataset on breast cancer survival in South Africa's public sector to date.

Conclusion

Socioeconomic disadvantage, late-stage diagnosis, and HIV co-infection are the strongest predictors of 5-year overall survival among South African women with breast cancer treated in the public sector. These determinants reflect both health system and upstream social inequities. Efforts to improve survival should integrate community-based interventions to address delayed diagnosis and poor awareness, while health systems must evolve to provide responsive care to high-risk populations, particularly for women living with HIV.

Acknowledgments

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy, or views of these organizations.

Author contributions

Maureen M. Joffe (Conceptualization, Funding acquisition, Writing—original draft), Wenlong Carl Chen (Formal analysis, Validation, Writing—review & editing), Ashleigh Craig (Formal analysis, Writing—review & editing), Daniel S. O'Neil (Writing—review & editing), Alfred I. Neugut (Funding acquisition, Writing—review & editing), Judith Jacobson (Conceptualization, Funding acquisition, Writing—review & editing), Paul Ruff (Conceptualization, Funding acquisition, Writing—review & editing), Rofhiwa M. Mathiba (Data curation, Investigation, Resources, Writing—review & editing), Nivashini M. Murugan (Data curation, Investigation, Resources, Writing—review & editing), Herbert Cubasch (Data curation, Funding acquisition, Investigation, Resources, Writing—review & editing), Charmaine M. Blanchard (Investigation, Resources, Writing—original draft), Sarah M. Nietz (Investigation, Resources, Writing—review & editing), Jennifer

M. Edge (Data curation, Investigation, Resources, Writing—review & editing), Ines Buccimazza (Data curation, Investigation, Resources, Writing—review & editing), Sharon Cacala (Data curation, Investigation, Resources, Writing—review & editing), Valerie McCormack (Conceptualization, Formal analysis, Writing—review & editing), Yoanna Pumpalova (Methodology, Writing—review & editing), and Shane Norris (Conceptualization, Formal analysis, Supervision, Writing—review & editing)

Supplementary material

Supplementary material is available at *The Oncologist* online.

Funding

This study was supported by the NIH/NCI grant; grant/award numbers: NIH/R01-CA19262701 and R01-CA250012; South African Medical Research Council/University of the Witwatersrand Common Epithelial Cancer Research Center (MRC/WITS CECRC); German Federal Ministry of Research, Technology and Space (BMFTR) 01KA2220B to the RHISSA Programme for the NORA Consortium; Science for Africa Foundation to the Programme Del-22-008 with support from Wellcome Trust and the UK Foreign, Commonwealth & Development Office and is part of the EDCPT2 programme supported by the European Union.

Conflicts of interest

A.N.: Otsuka, United Biosource Corp, Hospira, Value Analytics, Merck, Organon, and GlaxoSmithKline (Consulting/advisory relationship); EHE Intl (Scientific Advisory Board); Otsuka (Research Funding); P.R.: Merck, Roche, Pfizer, GSK, Jansen, AstraZeneca, ImmunityBio, Amgen (clinical trial funding to the institution). The other authors indicated no financial relationships related to this study.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics

The SABCHO Study was approved by the University of the Witwatersrand Human Research Ethics Committee (Medical) (M1911203), the University of KwaZulu-Natal Biomedical Research Committee (BF080/15), and the Institutional Review Board of Columbia University (IRB-AAAQ135).

References

1. Arnold M, Morgan E, Rumgay H, et al. Current and future burden of breast cancer: global statistics for 2020 and 2040. *Breast*. 2022;66:15-23. https://doi.org/10.1016/j.breast.2022.08.010
2. Lei S, Zheng R, Zhang S, et al. Global patterns of breast cancer incidence and mortality: a population-based cancer registry data analysis from 2000 to 2020. *Cancer Commun (Lond)*. 2021;41:1183-1194. https://doi.org/10.1002/cac2.12207
3. Kim J, Harper A, McCormack V, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. *Nat Med*. 2025;31:1154-1162. https://doi.org/10.1038/s41591-025-03502-3
4. Rezaeian S, Khazaei S, Khazaei S, Mansori K, Sanjari Moghaddam A, Ayubi E. Human development inequality index and cancer pattern: a global distributive study. *Asian Pac J Cancer Prev*. 2016;17:201-204. https://doi.org/10.7314/apjcp.2016.17.s3.201
5. Igbokwe KK. Comparative examination of breast cancer burden in Sub-Saharan Africa, 1990-2019: estimates from Global Burden of Disease 2019 study. *BMJ Open*. 2024;14:e082492. https://doi.org/10.1136/bmjopen-2023-082492
6. Limenih MA, Mekonnen EG, Birhanu F, et al. Survival patterns among patients with breast cancer in Sub-Saharan Africa: a systematic review and meta-analysis. *JAMA Netw Open*. 2024;7:e2410260. https://doi.org/101001/jamanetworkopen202410260
7. McCormack V, McKenzie F, Foerster M, et al. Breast cancer survival and survival gap apportionment in Sub-Saharan Africa (ABC-DO): a prospective cohort study. *Lancet Glob Health*. 2020;8:e1203-e1212. https://doi.org/10.1016/s2214-109x(20)30261-8
8. Scanlon B, Brough M, Wyld D, Durham J. Equity across the cancer care continuum for culturally and linguistically diverse migrants living in Australia: a scoping review. *Global Health*. 2021;17:87. https://doi.org/10.1186/s12992-021-00737-w
9. Joffe M, Ayeni O, Norris SA, et al. Barriers to early presentation of breast cancer among women in Soweto, South Africa. *PLoS One*. 2018;13:e0192071-e0192071. https://doi.org/10.1371/journal.pone.0192071
10. McKenzie F, Zietsman A, Galukande M, et al. Drivers of advanced stage at breast cancer diagnosis in the multicountry African breast cancer—disparities in outcomes (ABC-DO) study. *Int J Cancer*. 2018;142:1568-1579. https://doi.org/10.1002/ijc.31187
11. Gbenonsi GY, Martini J, Mahieu C. An analytical framework for breast cancer public policies in Sub-Saharan Africa: results from a comprehensive literature review and an adapted policy Delphi. *BMC Public Health*. 2024;24:1535. https://doi.org/10.1186/s12889-024-18937-5
12. Mapanga W, Norris SA, Craig A, et al. Drivers of disparities in stage at diagnosis among women with breast cancer: South African breast cancers and HIV outcomes cohort. *PLoS One*. 2023;18:e0281916. https://doi.org/10.1371/journal.pone.0281916
13. Leung Soo C, Pant Pai N, Bartlett SJ, Esmail A, Dheda K, Bhatnagar S. Socioeconomic factors impact the risk of HIV acquisition in the township population of South Africa: a Bayesian analysis. *PLOS Glob Public Health*. 2023;3:e0001502. https://doi.org/10.1371/journal.pgph.0001502
14. Chhatre S, Schapira M, Metzger DS, Jayadevappa R. Association between HIV infection and outcomes of care among Medicare enrollees with breast cancer. *EClinicalMedicine*. 2019;17:100205. https://doi.org/10.1016/j.eclinm.2019.11.001
15. Francis D, Webster E. Inequality in South Africa. *Dev South Afr*. 2019;36:733-734. https://doi.org/10.1080/0376835X.2019.1699397
16. Mayosi BM, Benatar SR. Health and health care in South Africa—20 years after Mandela. *N Engl J Med*. 2014;371:1344-1353. https://doi.org/10.1056/NEJMsr1405012
17. Cubasch H, Ruff P, Joffe M, et al. South African Breast Cancer and HIV Outcomes study: methods and baseline assessment. *J Glob Oncol*. 2017;3:114-124. https://doi.org/10.1200/JGO.2015.002675
18. Engel GL. The clinical application of the biopsychosocial model. *Am J Psychiatry*. 1980;137:535-544. https://doi.org/10.1176/ajp.137.5.535
19. Mapanga W, Ayeni OA, Chen WC, et al. The South African breast cancer and HIV outcomes study: profiling the cancer centres and cohort characteristics, diagnostic pathways, and treatment approaches. *PLOS Glob Public Health*. 2023;3:e0002432. https://doi.org/10.1371/journal.pgph.0002432
20. Ayeni OA, Joffe M, Mapanga W, et al. Multimorbidity and overall survival among women with breast cancer: results from the South

African Breast Cancer and HIV Outcomes study. *Breast Cancer Res*. 2023;25:7. https://doi.org/10.1186/s13058-023-01603-w
21. Rachet B, Maringe C, Woods LM, Ellis L, Spika D, Allemani C. Multivariable flexible modelling for estimating complete, smoothed life tables for sub-national populations. *BMC Public Health*. 2015;15:1240. https://doi.org/10.1186/s12889-015-2534-3
22. Liu Q, Charleston MA, Richards SA, Holland BR. Performance of Akaike Information Criterion and Bayesian Information Criterion in selecting partition models and mixture models. *Syst Biol*. 2023;72:92-105. https://doi.org/10.1093/sysbio/syac081
23. Caswell-Jin JL, Sun LP, Munoz D, et al. Analysis of breast cancer mortality in the US-1975 to 2019. *JAMA*. 2024;331:233-241. https://doi.org/10.1001/jama.2023.25881
24. Coghill AE, Shiels MS, Suneja G, Engels EA. Elevated cancer-specific mortality among HIV-infected patients in the United States. *J Clin Oncol*. 2015;33:2376-2383. https://doi.org/10.1200/JCO.2014.59.5967
25. Foy KC, Fisher JL, Lustberg MB, Gray DM, DeGraffinreid CR, Paskett ED. Disparities in breast cancer tumor characteristics, treatment, time to treatment, and survival probability among African American and white women. *NPJ Breast Cancer*. 2018;4:7. https://doi.org/10.1038/s41523-018-0059-5
26. Wang S, Tang W, Wang S, Hong S, Liu J. Racial disparities in survival of breast cancer patients after surgery. *Front Public Health*. 2022;10:831906. https://doi.org/10.3389/fpubh.2022.831906
27. Chasimpha S, McCormack V, Cubasch H, et al. Disparities in breast cancer survival between women with and without HIV across Sub-Saharan Africa (ABC-DO): a prospective, cohort study. *Lancet HIV*. 2022;9:e160-e171. https://doi.org/10.1016/S2352-3018(21)00326-X
28. Gnangnon FHR, Parente A, Aboubakar M, et al. Prognostic factors and overall survival of breast cancer in Benin: a hospital-based study. *BMC Womens Health*. 2024;24:295. https://doi.org/10.1186/s12905-024-03114-y
29. Songiso M, Nunez O, Cabanes A, et al. Three-year survival of breast cancer patients attending a one-stop breast care clinic nested within a primary care health facility in Sub-Saharan Africa-Zambia. *Int J Cancer*. 2024;155:261-269. https://doi.org/10.1002/ijc.34920
30. Coughlin SS. Social determinants of breast cancer risk, stage, and survival. *Breast Cancer Res Treat*. 2019;177:537-548. https://doi.org/10.1007/s10549-019-05340-7
31. Councell KA, Polcari AM, Nordgren R, Skolarus TA, Benjamin AJ, Shubeck SP. Social vulnerability is associated with advanced breast cancer presentation and all-cause mortality: a retrospective cohort study. *Breast Cancer Res*. 2024;26:176. https://doi.org/10.1186/s13058-024-01930-6
32. WHO. Global cancer burden growing, amidst mounting need for services. 2024. Accessed August 8, 2024. https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing—amidst-mounting-need-for-services