Summer 2025 | Volume 52 | Number 2

Journal of Registry Management

Published by the National Cancer Registrars Association
Founded in 1975 as *The Abstract*

ISSN: 1945-6123

The Right Job for the Ideal Candidate

JOB BANK IS NCRA'S RESOURCE CONNECTING EMPLOYERS TO CANCER REGISTRARS.

OPPORTUNITY AWAITS!

Visit Job Bank today at
www.ncra-usa.org/jobs.

Journal of Registry Management

is published quarterly by the
National Cancer Registrars Association
1330 Braddock Place, #520
Alexandria, VA 22314
(703) 299-6640
(703) 299-6620 FAX

Address any correspondence to:
National Cancer Registrars Association
1330 Braddock Place, #520
Alexandria, VA 22314

Address all editorial correspondence to:
Nadine Walker, MS, ODS-C
Email: JRMEditor@ncra-usa.org
Telephone: (703) 299-6640 ext. 327

Letters to the Editor
Letters to the Editor must be signed and include address and telephone number; none will be published anonymously. Letters subject to editing.

Editorial Policy
Opinions expressed in articles appearing in *Journal of Registry Management* are those of the authors and not necessarily those of the National Cancer Registrars Association, or the editor of *Journal of Registry Management*.

Accessing the *JRM*
Access to the *JRM* is available via NCRA's website at the *JRM* Library (ncra-usa.org). Full-text articles are also available on the National Library of Medicine's PubMed Central website.

Advertising Sales
Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the *Journal* does not indicate or imply endorsement by the *Journal* or NCRA. Advertising sales should be directed to Marketing Coordinator, NCRA, 1330 Braddock Place, #520, Alexandria, VA 22314, (703) 299-6640, (703) 299-6620 FAX.

National Cancer Registrars Association

Journal of Registry Management

Summer 2025 • Volume 52 • Number 2

Contents

Indexing

The *Journal of Registry Management* is indexed in the National Library of Medicine's MEDLINE database. Citations from the articles indexed, the indexing terms (key words), and the full article are included and searchable using PubMed.

For your convenience, the *Journal of Registry Management* is indexed in the 4th issue of each year and on the Web (under "Resources" at http://www.ncra-usa.org/jrm). The 4th issue indexes all articles for that particular year. The Web index is a cumulative index of all *JRM* articles ever published.

Letter from the Editor

Dear Readers,

This issue of the *JRM* supports the belief that cancer registries are a core resource within cancer surveillance systems globally and provide valuable information that informs cancer treatment and prevention efforts worldwide. Inside is an original article that summarizes a study looking at the documentation of HIV status for people diagnosed with cancer in one of eastern Africa's oldest population-based cancer registries. There is also original research that explores disparities by sex and race that exist within Arkansas through the lens of that state's obesity and cancer burden. Another original article suggests an approach to reduce missingness for race, a core demographic component of cancer surveillance. This issue also contains 7 Success Story posters from the CDC's National Program of Cancer Registries (NPCR) that highlight the importance and significance of cancer data reporting and timeliness in the U.S.

"*With advances in care for both HIV and cancers, countries with high burdens, especially those in Africa, have established cancer registries and HIV clinical care databases.*" Pallangyo and coauthors assessed how often HIV status was recorded in the cancer registry database, and looked at its documentation in the health record, highlighting considerations for self-reporting HIV status by gender. The study also emphasized which malignancies had HIV status recorded most often, as well as the most common HIV-positive malignancies. Additionally, data linkages between cancer registries and HIV databases were determined to be lacking.

Daniela Ramirez Aguilar and coauthors explored obesity-associated cancer rates among adults in Arkansas from 2010–2019, utilizing that state's cancer registry data guided by the CDC's definitions of obesity-associated cancer. Their analysis provided an evaluation looking at sex and race and determined that Black women in Arkansas had the highest rate of obesity-associated cancers, and that among all groups, colorectal cancer was the most frequently seen obesity-associated cancer. The study also found that White men who were obese had a much higher rate of esophageal cancer. Cancer registry data with race as a factor is also the focus of the article by Francis Boscoe, who describes a methodological approach, Bayesian Improved Surname Geocoding (BISG), that utilizes patient surnames and addresses to predict race and ethnicity. The article suggests that this approach, which is available for use as a software package, could help improve race and ethnicity missingness in central cancer registries with consideration of the sensitivity of the results.

The CDC-NPCR Success Story posters from several state registries also illustrate how they support data reporting and data timeliness. The posters were on display at the National Cancer Registrars Association annual educational conference from May 3–6, 2025, and are from Arizona, Louisiana, Maine, Maryland, Mississippi, Ohio, and Oklahoma. The posters provide a snapshot of efforts to enhance and improve cancer surveillance data.

Finally, this issue's quiz is derived from the article, *Inferring Unknown Race in Central Cancer Registries*. The quiz was submitted by *JRM* contributing editor Cari Vida, RHIA, ODS-C, and offers the opportunity to not only test your knowledge but also earn continuing education (CE) credits.

As a reminder, you can access previously published articles at **https://www.ncra-usa.org/JRM**, and you can also find the *JRM* on **PubMed (nih.gov)**.

On behalf of the editorial and production teams and the editorial advisory board, we extend thanks to every author, contributor, and reader for supporting the *JRM*.

Nadine R. Walker, MS, ODS-C
JRM Editor-in-Chief

Original Article

Burden of HIV among Patients Undergoing Cancer Treatment: Analysis of Population Cancer Registry in Northern Tanzania

Angela Pallangyo[a,b,c]; Onstard Mashauri[a]; Salum Kalonge[c]; Maryam Amour[d]; Alex Mremi[a,b,c]; James S. Ngocho[a]; Emmanuel Balandya[d]; Gideon Kwesigabo[d]; Benson Kidenya[e]; Stephen E. Mshana[e]; Eligius F. Lyamuya[d]; Bruno F. Sunguya[d]; John Bartlett[f]; Blandina T. Mmbaga[a,b,c]

Abstract: Background: The burden of cancer in sub-Saharan African countries is escalating with a rising Human Immunodeficiency Virus (HIV) prevalence. However, information about the burden of HIV on cancer epidemiology is scarce. Specifically, little is known about HIV infection among the cancer cases registered in the Kilimanjaro Population Cancer Registry (KCR) despite the presence of this infection in Tanzania. Thus, our study aimed to assess the burden of HIV in cancer patients by evaluating HIV serostatus information among recorded cases of malignancies in the KCR. Methods: This secondary data analysis examined records of all cancer cases registered in the KCR from January 2018 through December 2022 to assess the status of HIV infection among the cancer cases. Variables assessed were demographic information, type of cancer, and HIV serostatus. Proportions were analyzed using descriptive data. Results: A total of 5,508 cancer cases were recorded from 2018 through 2022. HIV serostatus was documented in 4.8% (226/5,508) of the cancer cases, 68% of which were HIV seropositive with a slight female predominance (male-to-female ratio of 1:1.7). Cervical cancer was the leading malignancy (18%) with recorded HIV serostatus. Patients aged 18–50 years and females had the highest prevalence of HIV infection (64.6% and 63.5%). Conclusion: HIV infection is still underreported among cancer patients in the cancer registry of the Kilimanjaro region with only 4.8% of malignancies registered in KCR having a documented HIV serostatus. HIV serostatus was mostly documented in AIDS-defining cancers. Thus, efforts to support HIV counseling and testing among cancer patients should be made, as this will also affect treatment plans and monitoring.

Key words: *Cancer registry, HIV, Tanzania*

Introduction

Cancer is an emerging public health problem in developing countries. The burden of cancer is rapidly increasing in sub-Saharan African (SSA) countries, with an incidence of 801,392 (4.2% of 19.3 million cases) cases and mortality of about 520,158 (5.2% of 9.9 million cases) people in 2020,[1] which is a 2/5 increase since 2008. HIV infection has been concomitantly associated with this steep rise.[2]

By compromising the body's immunity, HIV infection is associated with disease progression in several cancers.[3] Some cancers have been associated with HIV infection.[3] These include Kaposi sarcoma and non-Hodgkin lymphoma in people living with HIV (PLWHIV). These cancers are referred to as AIDS-defining cancers (ADC), while those not linked with HIV, such as liver cancer, anal cancer, and breast cancer, are termed non-AIDS-defining cancers (NADC). With the use of antiretroviral therapy (ART), NADCs are becoming important causes of morbidity and mortality.[4]

With advances in care for both HIV and cancers, countries with high burdens, especially those in Africa, have established cancer registries and HIV clinical care databases. Tanzania is no exception.[5] These cancer registries and HIV clinical care databases may be interlinked in advanced health systems, making identification, diagnosis, and care easier and more effective. In higher-income countries, this has resulted in improved clinical outcomes and survival. This contrasts with the situation in low- and middle-income countries (LMICs), including Tanzania. In these contexts, patients suffering from both HIV and cancer are more likely to have their cancer identified and treated than HIV. Their HIV status is less likely to be recognized and thus, less likely to be treated.[4]

Tanzania has 4 cancer registries under the African Cancer Registry Network (AFCRN),[6] one of which is the Kilimanjaro Population Cancer Registry (KCR), one of the oldest population-based cancer registries, which was established in 1998 and covers about 1.5 million people.[7] The Kilimanjaro Christian Medical Centre (KCMC) Population Cancer Registry joined AFCRN in 2017. However, a representative association between cancers and HIV infection has not been properly documented in northern Tanzania, which is especially important in the era of free ART, which

[a]KCMC University, Moshi, Tanzania. [b]Kilimanjaro Christian Medical Centre, Moshi, Tanzania. [c]Kilimanjaro Clinical Research Institute, Moshi, Tanzania. [d]Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania. [e]Catholic University of Health and Allied Sciences, Mwanza, Tanzania. [f]Duke University, Durham, North Carolina.

Address correspondence to Angela Pallangyo, School of Medicine, KCMC University, PO Box 2240, Moshi, Longuo B, Kilimanjaro, Tanzania. Email: angela.pallangyo@kcmuco.ac.tz.

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

This study was funded by the Fogarty International Center of the National Institutes of Health under Award Number R25 TW011227, in an award that was won by the Transforming Health Professions Education in Tanzania (THET) consortium.

prolongs the lifespan of PLWHIV. Thus, our project aimed to evaluate the reporting of HIV status among cancer patients documented in KCR.

Methods

This cross-sectional study was conducted at the Cancer Care Centre (CCC) at the KCMC Hospital in the northern zone. The CCC caters to patients from the Tanga, Kilimanjaro, Manyara, and Arusha regions, which comprise more than 15 million people.[8] The study was conducted within the Kilimanjaro Population Cancer Registry (KCR), which has been housed under the Oncology Department of the Hospital since 2017. Cancer data was collected passively and actively within the registry. The registry data was collected from many sources, such as hospital medical records, pathology laboratories, and radiology departments, and included demographic, clinical, treatment, and follow-up information.

Medical records were the major source of cancer data in the cancer registry, while pathology reports were also used to confirm the diagnosis. HIV status was collected from the medical history of patients at the oncology department from January 2018 through December 2022.

Data was de-identified, cleaned, and analyzed using the statistical software Stata 15. Descriptive analysis was done by summarizing categorical variables as frequencies and percentages. The main categorical variable evaluated was HIV status. Other categorical variables assessed in this study included sex, age group, year of diagnosis, and type of cancer. The serostatus (positive or negative) among cancer patients with documented HIV information was calculated as a proportion. The association between HIV status and other variables could not be established due to the low proportion of HIV status documentation among cancer patients.

Results

A total of 5,508 confirmed cases of malignancies were recorded in the KCR during the 5 years of this study period. Among 5,508 confirmed cancer cases in the cancer registry, only 4.8% (266) of the patients had documented HIV status, and 95.2% (5,242) of patients had unknown HIV status, as shown in Figure 1.

Figure 1. Human Immunodeficiency Virus (HIV) Status Documentation in the Cancer Records at Kilimanjaro Population Cancer Registry (KCR)

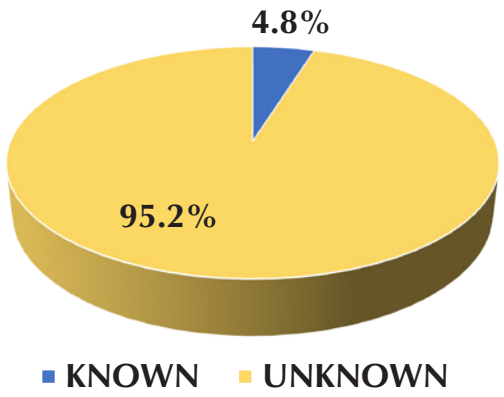
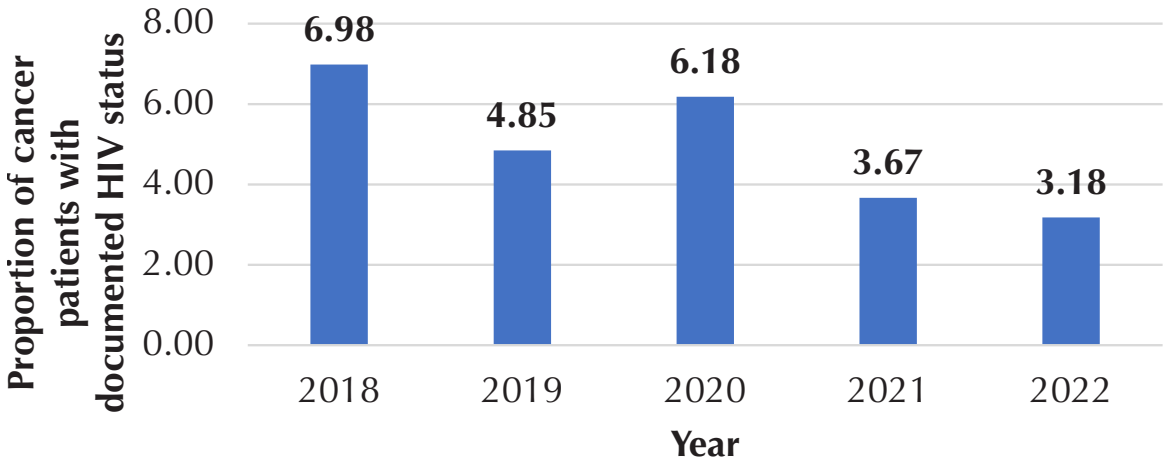


Figure 2. Proportion of Cancer Patients with Documented Human Immunodeficiency Virus (HIV) Status by Year



The proportion of cancer patients with documented HIV status declined over the 5 years from 2018 to 2022. In 2018, 6.98% of cancer patients had documented HIV status, a proportion which had marginally declined, to 4.85%, by 2019. The trend rose to 6.18% in 2020, but then decreased regularly, reaching 3.67% in 2021 and 3.18% in 2022 (Figure 2). The leading sex and age group with documented HIV status in the KCR during the study period were females (54.9%) and persons aged 18–50 years (51.9%). These findings are based on cancer reports that included documentation of HIV status.

Cervical malignancy (18.94%), Kaposi sarcoma (17.42%), lymphoma (11.36%), prostate (7.2%), and breast (6.82%) cancers were among the top 5 cancers of patients with recorded HIV status (Figure 3).

The proportion of documented HIV-positive status was substantially greater among females (72.8%) than males (61.1%). HIV serostatus was more prevalent (84.8%) in patients aged 18–50 years compared to other age groups. In terms of types of malignancies, Kaposi sarcoma had the

highest prevalence of HIV infection, with 93.5% of patients testing positive. Meanwhile, 80% of cervical cancer patients and 63.3% of lymphoma patients were seropositive for HIV. This is summarized in Table 1.

HIV was most prevalent among males with Kaposi sarcoma and among women with cervical cancer, as highlighted in the bar charts (Figure 4).

Discussion

Our study found that only 4.8% of malignancies recorded in the KCR were found in patients with a known HIV status. Furthermore, HIV serostatus was more frequently reported among patients with cervical cancer, Kaposi sarcoma, and lymphomas. HIV-positive cancer patients were most likely to have hematologic (Kaposi sarcoma), cervical, or lymphoid cancers. The age group of 18–50 years and female sex were found to have greater prevalence of HIV infection than others.

Our study found a low rate of recording the HIV status of cancer clients. This proportion is lower than those reported in other sub-Saharan African countries, such as Malawi, Zimbabwe, and South Africa, which have documented HIV infection status among 22%, 86%, and 92% of cancer patients, respectively, in their cancer registries.[9]

The CanReg 5 is the tool used for the population-based cancer registry and can capture the HIV status of cancer patients. Despite the tool's capacity to record HIV status, however, we still found a low proportion of cancer cases with documented HIV infection status. This low proportion might be due to poor recording of HIV status in each cancer patient's file rather than to lack of testing,[10] HIV serostatus not being routinely documented by clinicians in the investigation forms,[11] or biased documentation, whereby only AIDS or infection-related cancers trigger the clinician to investigate and document the serostatus of patients with cancer.[11] It is therefore recommended that HIV and cancer care services should be integrated into a single clinic to capture patients' HIV status.[10] Efforts should be made to ensure that HIV testing and care are initiated before the patient's cancer treatment.[10]

Figure 3. Proportion of Cancer Patients with Documented Human Immunodeficiency Virus (HIV) Status by Type of Cancer

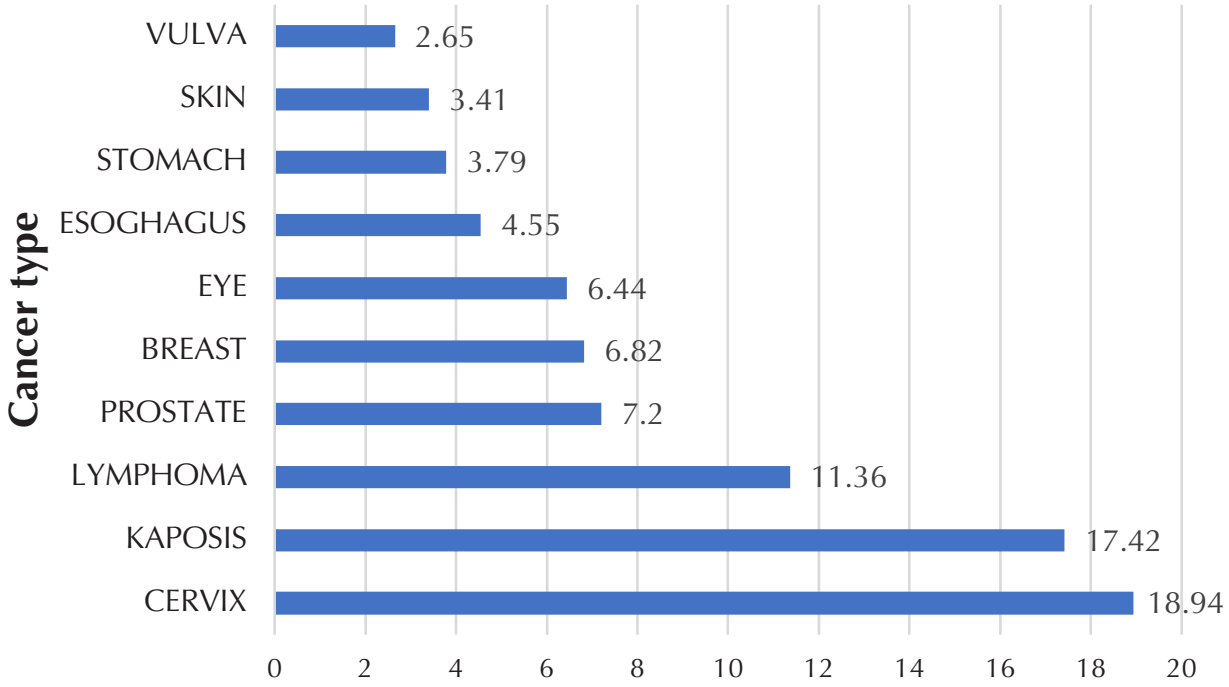
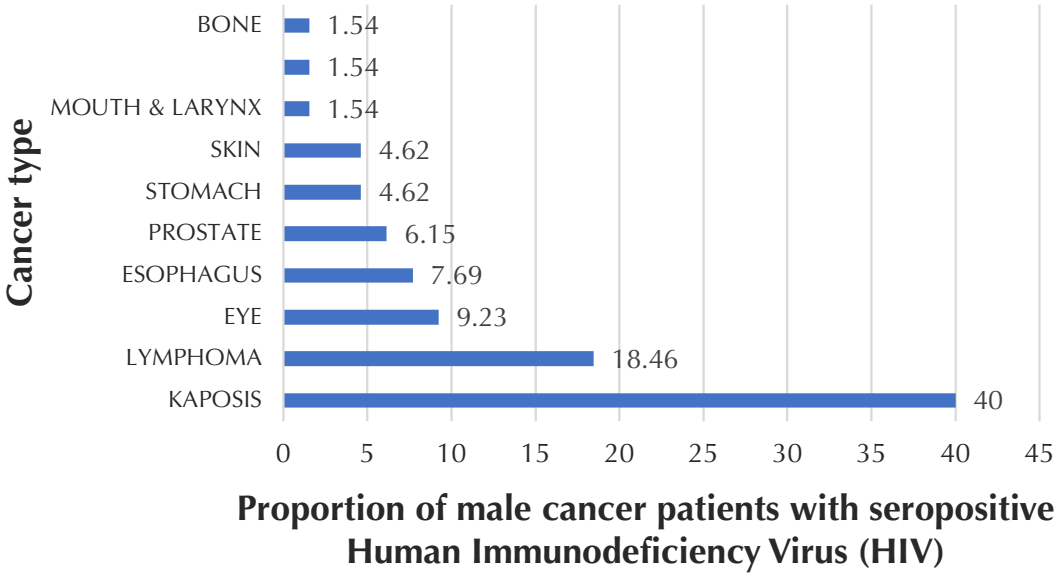
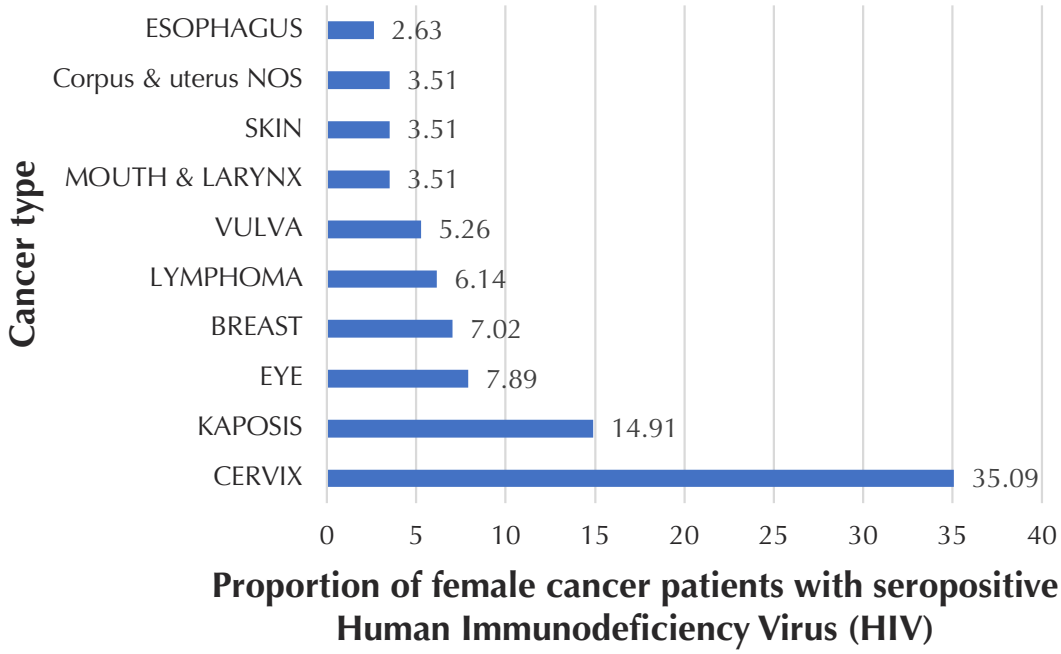


Table 1. Gender, Age Group, and Top 5 Cancers with Documented Serostatus of Cancer Patients and Corresponding Human Immunodeficiency Virus (HIV) Serostatus

Variable	*Serostatus*	
	Negative (n = 85)	**Positive (n = 181)**
Sex		
Male	42 (38.9)	66 (61.1)
Female	43 (27.2)	115 (72.8)
Age, y		
<18	5 (55.6)	4 (44.4)
18–50	21 (15.2)	117 (84.8)
>50	59 (49.6)	60 (50.4)
Cancer type (the top 5 cancers with documented HIV status)		
Cervix	10 (20.0)	40 (80.0)
Kaposi sarcoma	3 (6.5)	43 (93.5)
Lymphoma	11 (36.7)	30 (63.3)
Prostate	15 (78.9)	4 (21.1)
Breast	10 (55.6)	8 (44.4)

Figure 4 A. Top 10 Cancers of Males and B: Top 10 Cancers of Females Against the Proportion of Human Immunodeficiency Virus (HIV) Infection





There is no linkage between the cancer registry and HIV clinics in our country, which is contrary to common practice in other East African countries[12] and in Western countries, which have a well-formed data linkage between these 2 diseases. This is due to the lack of synchronization of private health facilities' medical data and laboratory services with that of public hospitals and laboratories at the local and national level.[11]

The most common malignancies recorded among the HIV-positive patients were Kaposi sarcoma, cervical cancer, and lymphomas.[14] Furthermore, these are the most common AIDS-defining cancers in Tanzania.[8] These findings are similar to those of other studies done in Tanzania[13] and similar to findings in developed countries, even though Western countries have a high prevalence of non-AIDS-defining cancers, such as breast cancers.[14] This may be due to wide availability of antiretrovirals that improve the lifespan of PLWHIV in developed countries compared to lower availability in low- and middle-income countries like Tanzania.[13,15]

Our study found that adults aged 18–50 years and females were the most likely to be affected by HIV infection. These findings are similar to those of studies done in sub-Saharan Africa, which have also found that adults in this age group and females were the most likely to be affected by HIV infection.16 However, these research findings are contrary to what is seen in the developed world, where HIV infection is most prevalent in adult

males.[17] These differences in sex prevalence could be due to different risk exposures in the 2 populations, as homosexuality is more common in developed countries[17] than it is in sub-Saharan Africa, where females also face different obstacles to exposing their HIV status, such as health care issues, including stigma by health care personnel, which is worsened by the fact that HIV-positive patients must carry their HIV care card with them to receive health services. Furthermore, violence from male partners, abandonment, and stigma from the community reduce the likelihood of women disclosing their HIV status.[10,18] Several mechanisms help improve the reporting of HIV status among women. These include tailored programs, such as nurse-facilitated disclosure among women and their partners; pre- and post-HIV testing counselling, which helps women understand the importance and implications of disclosure of HIV diagnosis; close collaboration between seropositive women and their families, treatment supporters, and primary health care providers to optimize the outcome; and peer group support groups for HIV-infected women. However, health care personnel need ongoing training in patient-centered counselling, to improve their awareness of and respect for multiple factors that may affect disclosure, including, but not limited to, sex, age, literacy, and social influence.[19,20]

Conclusion

There is still an underreporting of HIV infection among cancer patients in the KCR, where the cancer registry is still not linked with the HIV clinic database. To overcome this challenge, we recommend that HIV and cancer management clinics be integrated, that HIV testing and treatment should begin for every cancer patient before the initiation of cancer management, and that care and treatment clinics, hospital laboratories, and the cancer registry should be linked together. This will facilitate effective prevention and follow-up with cancer patients infected by HIV. Furthermore, we recommend interventional research on the topic of cancer registry and HIV documentation involving a larger sample size.

Acknowledgments

I am grateful to the AFCRN and the Network for Oncology Research in Sub-Saharan Africa (NORA) project, which support the Kilimanjaro Population Cancer Registry, and to KCMC Hospital, through its oncology departments, which allowed me to access patients' data.

I offer my sincere appreciation to my senior leaders, mentors, and fellow mentees for their excellent guidance, patience, encouragement, and support, which have helped to make this report real.

References

1. Bray F, Parkin DM, Gnangnon F, et al. Cancer in sub-Saharan Africa in 2020: a review of current estimates of the national burden, data gaps, and future needs. *Lancet Oncol.* 2022;23(6):719-728. doi:10.1016/S1470-2045(22)00270-4
2. Harlan JM, et al. Association between HIV infection and cancer stage at presentation at the Uganda Cancer Institute. *J Glob Oncol.* 2018:4:1-9. 2023:1-9. doi:10.1200/JGO.17.00005
3. 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(21):2376-2383. doi:10.1200/JCO.2014.59.5967
4. Godbole SV, Nandy K, Gauniyal M, et al. HIV and cancer registry linkage identifies a substantial burden of cancers in persons with HIV in India. *Medicine (Baltimore).* 2016;95(37):e4850. doi:10.1097/MD.0000000000004850
5. Anwar W, Boulos K, Chan M, African M. Cancer registries in Africa : An emerging need. *J Afr. Cancer.* 2012;4:140-141. doi:10.1007/s12558-012-0231-z
6. Seraphin TP, Joko-Fru WY, Kamaté B, et al. Rising prostate cancer incidence in sub-Saharan Africa: a trend analysis of data from the African Cancer Registry Network. *Cancer Epidemiol Biomarkers Prev.* 2021;30(1):158-165. doi:10.1158/1055-9965.EPI-20-1005
7. Samoyo PTK, Nkya GZ, Minja FG, Temu RJ. Clinicopathological guide to malignant bone tumours: A retrospective analysis of the cancer registry at Kilimanjaro Christian Medical Centre in northern Tanzania. *East Cent Afr J Surg.* 2018;22(2):24. doi:10.4314/ecajs.v22i2.4
8. Mremi A, Mswima J, Mlay MG, et al. Cancer spectrum in HIV-infected patients: a zonal hospital experience in Tanzania. *Cancer Treat Res Commun.* 2020;25:100213. doi:10.1016/j.ctarc.2020.100213
9. Africa S, Montaño MA, Mtisi T, et al. Characterizing HIV status documentation among cancer patients at regional cancer centers in Malawi, Zimbabwe, and South Africa. *medRxiv [preprint].* Published online September 5, 2023. doi:10.1101/2023.09.04.23294963
10. Montaño MA, Mtisi T, Ndlovu N, et al. Characterizing HIV status documentation among cancer patients at regional cancer centers in Malawi, Zimbabwe, and South Africa. *Infect Agent Cancer.* 2023;18(1):65. doi:10.1186/s13027-023-00548-1
11. Dhokotera T, Bohlius J, Spoerri A, et al. The burden of cancers associated with HIV in the South African public health sector, 2004–2014: A record linkage study. *Infect Agent Cancer.* 2019;14(1). doi:10.1186/s13027-019-0228-7
12. Mbulaiteye SM, Katabira ET, Wabinga H, et al. Spectrum of cancers among HIV-infected persons in Africa: The Uganda AIDS-Cancer Registry Match Study. *Int J Cancer.* 2006;118(4):985-990. doi:10.1002/ijc.21443
13. Chalya PL, Mbunda F, Rambau PF, et al. Kaposi's sarcoma: A 10-year experience with 248 patients at a single tertiary care hospital in Tanzania Cancer. *BMC Res Notes.* 2015;8:440. doi:10.1186/s13104-015-1348-9
14. Mathoma A, Sartorius B, Mahomed S. The trends and risk factors of AIDS-defining cancers and non-AIDS-defining cancers in adults living with and without HIV: a narrative review. *J Cancer Epidemiol.* 2024;2024:1-16. doi:10.1155/2024/7588928
15. Cortes J, Perez-García JM, Llombart-Cussac A, et al. Enhancing global access to cancer medicines. *CA Cancer J Clin.* 2020;70(2):105-124. doi:10.3322/caac.21597
16. Steenkamp L, Venter D, Walsh C, Dana P. Socio-economic and demographic factors related to HIV status in urban informal settlements in the Eastern Cape, South Africa. *Afr J AIDS Res.* 2014;13(3):271-279. doi:10.2989/16085906.2014.952651
17. Woodring J, Kruszon-Moran D, McQuillan G. HIV infection in U.S. household population aged 18–59: data from the National Health and Nutrition Examination Survey, 2007–2012. *Natl Health Stat Report.* 2015;(83):1-13.
18. Loutfy M, Johnson M, Walmsley S, et al. The association between HIV disclosure status and perceived barriers to care faced by women living with HIV in Latin America, China, Central/Eastern Europe, and Western Europe/Canada. *AIDS Patient Care STDS.* 2016;30(9):435-444. doi:10.1089/apc.2016.0049
19. Geubbels E, Williams A, Ramaiya A, Tancredi D, Young S, Chantry C. HIV status disclosure among postpartum women in rural Tanzania: predictors, experiences and uptake of a nurse-facilitated disclosure intervention. *AIDS Care.* 2018;30(4):417-425. doi:10.1080/09540121.2018.1428724
20. Mwamba RN, Sao SS, Knettel BA, et al. The disclosure dilemma: willingness to disclose a positive HIV status among individuals preparing for HIV testing during antenatal care in Tanzania. *AIDS Behav.* 2021;25(3):908-916. doi:10.1007/s10461-020-03058-y

Original Article

Quantifying Cancer Burden Attributable to Obesity: Highlighting the Disparities by Sex, Race, and Ethnicity in a Rural State with High Obesity and Cancer Burden

Daniela Ramirez Aguilar, MPH;[a] Yong-Moon Park, MD, PhD;[b] Mario Schootman, PhD;[b] Jaimi L. Allen, PhD;[b] Michael R. Thomsen, PhD;[b] Nithya Neelakantan, PhD;[a,b] Bala Simon, MD, DrPH[a,b]

Abstract: Overweight and obesity is a complex, multifactorial disease that increases the risk of several cancers. The purpose of this study is to calculate the population-attributable fraction (PAF%) of obesity-associated cancer rates among adults from 2010–2019 in Arkansas by sex, race, and ethnicity. Obesity-associated cancer data for this period were obtained from the Arkansas Central Cancer Registry. The PAF% was calculated using obesity prevalence and global relative risks. Obesity prevalence data were gathered from the Arkansas Behavioral Risk Factor Surveillance System. Global relative risks for each obesity-associated cancer were gathered from large-scale epidemiological studies, meta-analyses, and systematic reviews in which body mass index (BMI) was ≥ 30 kg/m^2. Obesity-attributable cancer age-adjusted incidence rates (AAIRs) were calculated by multiplying the obesity-associated cancer's AAIR by the estimated PAF%. Breast, esophageal adenocarcinoma, gallbladder, kidney, and liver obesity-associated cancers each had a PAF% greater than 25% by sex, race, and ethnicity. Arkansas non-Hispanic (NH) Black women were disproportionately impacted by obesity-attributable cancers, with a disparity driven primarily by the higher incidence of breast and other female-specific cancers. Findings suggest that targeted screening among those with BMI ≥ 30 kg/m^2 could decrease the burden of obesity-associated and attributable cancers in Arkansas, particularly for breast and colorectal cancers.

Key words: *age-adjusted incidence rates, Arkansas, epidemiology, obesity-associated cancers, population attributable fraction*

Objective

Obesity is a modifiable risk factor for various chronic diseases.[1] It is a complex, multifactorial disease that increases the risk of several cancers.[2] The incidence of many of these cancers has been increasing, possibly due to the increasing prevalence of obesity. According to a Centers for Disease Control and Prevention (CDC) study, about 40% of all new cancer diagnoses in the United States are associated with overweight and obesity.[3] Of the 13 cancer types associated with obesity, at least 5 (breast, colorectal, corpus uteri, kidney, liver, pancreatic, and thyroid) are among the 10 most commonly diagnosed cancers.[4] While these observed rising cancer rates and obesity prevalence can only imply an association, continuing increases in obesity prevalence over time also suggest that the burden will likely increase in the decades to come.[5] Evidence has also shown an increase in the risk of cancer recurrence and mortality among cancer survivors with obesity, underscoring the need for early interventions for patients before, during, and after a cancer diagnosis.[6,7] Data are needed to identify populations for whom the burden is highest in a rural, racially diverse, and high-poverty southern state (Arkansas) with a high obesity and cancer burden.[8] Even though cancers associated with obesity quantify the relationship between exposure and disease (obesity-associated cancer), this association does not provide an estimate for cancer cases that are likely due to obesity (obesity-attributable cancers). Knowing the burden of cancers attributable to obesity can help clinicians assess the impact of obesity on cancer.[9] These results may help community outreach programs to stress the importance of obesity and cancer prevention efforts.

For this study, data were used from the Arkansas Central Cancer Registry, a population-based registry that identifies cancer trends and rates at a state level. By using central cancer registry data, this study describes the burden of obesity-associated cancers and estimates attributable rates utilizing clinically reported cancer cases. Understanding the burden of cancer attributable to obesity is crucial, especially for Arkansas, due to its racial and ethnic diversity and its complex socioeconomic elements, which include poverty, rurality, inadequate nutrition, the limited physical activity of its population, and their limited access to quality healthcare.[10,11] The purpose of this study is to calculate obesity-attributable cancer rates derived from the population attributable fraction of adults from 2010–2019 in Arkansas by sex, race, and ethnicity.

[a]Arkansas Department of Health, Little Rock, Arkansas. [b]University of Arkansas for Medical Sciences, Little Rock, Arkansas.

Address correspondence to Daniela Ramirez Aguilar, MPH, Arkansas Central Cancer Registry, Arkansas Department of Health, 4815 W. Markham Street, Slot 7, Little Rock, AR 7220. Email: Daniela.Ramirez-Aguilar@Arkansas.gov.

This publication was supported by DP22-2202 Cooperative Agreement #6 NU58DP007090-03-01 from the Centers for Disease Control and Prevention (CDC). Research support was partially provided by the National Heart, Lung, and Blood Institute (1K01HL175206, Park). The views expressed in this paper are solely those of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Arkansas Department of Health.

Methods

Obesity-Associated Cancer Rates

This cross-sectional study used data obtained from the Arkansas Central Cancer Registry for adult patients (≥ 18 years of age). Data were evaluated by sex and 2 major racial groups (non-Hispanic [NH] Black and White). Age-adjusted incidence rates (AAIR) for obesity-associated cancers were computed using SEER*Stat software version 8.4.3 with corresponding 95% CIs calculated as modified gamma intervals on a central cancer registry imported dataset.[12] Cancers were identified according to the CDC's definitions of obesity-associated cancers, which is a predefined selection available in SEER*Stat.[13] Cancers associated with obesity are defined by cancer type using an ICD-O-3 primary site and histology codes. Additional restrictions are applied to certain cancers: Esophageal adenocarcinoma (EAC) is restricted to cases microscopically confirmed, while corpus uteri (not otherwise specified) and ovarian cancer cases are restricted to women. Breast cancer is restricted to women who are over the age of 50 years and postmenopausal. For this analysis, meningioma (for all groups) and esophageal adenocarcinoma for (NH Black women) were excluded from analysis due to low case counts in Arkansas. Because the COVID-19 pandemic impacted cancer reporting for the diagnosis years 2020 and 2021, this study used prepandemic cancer diagnosis years 2010–2019. This study was determined to be exempt from Institutional Review Board review as it posed minimal risk to human subjects.

Population Attributable Fraction

The population attributable fraction (PAF) was calculated using published relative risks for each obesity-associated cancer and obesity prevalence by sex, race, and ethnicity group. Obesity prevalence for each group was obtained from the 2011 Arkansas Behavioral Risk Factor Surveillance System (BRFSS), as 2011 is the first year that the current BRFSS weighting methodology was implemented. Global relative risks for each obesity-associated cancer were gathered from large-scale epidemiological studies, meta-analyses, and systematic reviews where body mass index (BMI) was ≥ 30 kg/m^2 (Table 1). To calculate the obesity-attributable cancer rate, the AAIR of obesity-associated cancers was multiplied by the PAF (%)[9]:

Obesity-attributable cancer$_{AAIR}$ = obesity-associated cancer$_{AAIR}$ × PAF%

Table 1. Relative Risk for Obesity-Associated Cancers Gathered from Meta-analysis, Systematic Review, and Large-Scale Epidemiological Studies Where Obesity Has ≥ 30 Body Mass Index Classification

Obesity-Associated Cancer Type	*Relative Risk (95% CI)*
Breast, postmenopausal*	1.13 (1.05–1.22)[1]
Colorectal	1.33 (1.25–1.42)[2]
Corpus uteri*	3.22 (2.91–3.56)[1]
Esophageal adenocarcinoma	3.29 (1.82–5.95)[3]
Gallbladder	1.58 (1.43–1.75)[4]
Gastric cardia	1.13 (1.03–1.24)[5]
Kidney	1.76 (1.61–1.91)[6]
Liver	1.77 (1.56–2.01)[7]
Multiple myeloma	1.23 (0.99–1.52)[8]
Ovary*	1.28 (1.20–1.36)[1]
Pancreas	1.47 (1.23–1.75)[9]
Thyroid	1.09 (0.98–1.22)[10]

*Female-specific.

Results

Obesity-Associated Cancer Rates

In Arkansas, 58,698 obesity-associated cancers were diagnosed from 2010–2019. Overall, NH Black women had the highest rate of obesity-associated cancers (292.5 cases per 100,000 population). Among female-specific cancers, postmenopausal breast cancer had the highest incidence rate (NH Black women, 317.3 cases per 100,000 population; NH White women, 319.3 cases per 100,000 population). Among cancers that affect all groups, colorectal cancer had the highest frequency of cancer, with NH Black adults having the highest incidence rate (NH Black men = 73.9 cases per 100,000 population; NH Black women = 60.7 cases per 100,000 population) (Table 2).

Population Attributable Fraction (PAF) Estimates

The obesity crude prevalence in Arkansas was 59.9% (95% CI: 49.0-70.8) for NH Black women, 48.7% (95% CI, 44.2–53.3) for NH White women, 40.10% (95% CI, 29.2–51.0) for NH Black men, and 51.3% (95% CI, 46.7–55.8) for NH White men. Breast, esophageal adenocarcinoma, gallbladder, kidney, and liver obesity-associated cancers each accounted for at least one group with a PAF greater than 25%. Among female-specific cancers, NH Black women had a higher PAF than NH White women for obesity-attributable breast, corpus uteri, and ovarian cancer. Excluding NH Black women, the PAF for esophageal adenocarcinoma attributable to obesity among all groups ranged from approximately 47.9–54.0% (Table 3).

Estimated Obesity-Attributable Cancer Rates

Overall, NH Black women had a higher obesity-attributable cancer rate for breast cancer (36.9 cases per 100,000 population), colorectal cancer (10.1 cases per 100,000 population), cancer of the corpus uteri (18.8 cases per 100,000 population), gallbladder cancer (0.5 cases per 100,000 population), multiple myeloma (1.8 cases per 100,000 population), and pancreatic cancer (4.9 cases per 100,000 population). Among cancers that affect both men and women, NH White men had a higher obesity-attributable cancer rate for esophageal adenocarcinoma (3.67 cases per 100,000 population), gastric cardia (0.28 cases per 100,000 population), and kidney cancer (8.19 cases per 100,000 population) (Figure 1).

Discussion

Overall, publicly available data from 2010–2019 show that Arkansas and the United States have a similar obesity-associated rate of approximately 172 cancer cases per 100,000

Table 2. Number and Age-Adjusted Incidence Rates (AAIRs) of Obesity-Associated Cancers by Race/Ethnicity and Sex, Arkansas, 2010–2019

	Women				*Men*			
	Non-Hispanic Black		**Non-Hispanic White**		**Non-Hispanic Black**		**Non-Hispanic White**	
	Count	AAIR (95% CI)	Count	AAIR (95% CI)	Count	AAIR (95% CI)	Count	AAIR (95% CI)
Overall	5,624	292.5 (284.7–300.5)	34,544	266.0 (263.1–268.9)	2,415	166.0 (158.9–173.2)	16,115	143.7 (141.4–146.0)
Cancer Type								
Breast, postmenopausal*	2,233	317.3 (303.9–331.2)	15,490	319.3 (314.3–324.5)	-	-	-	-
Colorectal	1,118	60.7 (57.1–64.5)	5,927	45.7 (44.5–46.9)	1,070	73.9 (69.2–78.8)	6,556	59.0 (57.5–60.5)
Corpus uteri*	616	31.1 (28.6–33.7)	3,753	29.9 (29.0–31.0)	-	-	-	-
Esophageal adenocarcinoma	†	†	111	0.9 (0.7–1.0)	25	1.6 (1.0–2.4)	789	6.8 (6.3–7.3)
Gallbladder	35	1.9 (1.3–2.6)	185	1.4 (1.2–1.7)	16	1.2 (0.6–2.0)	85	0.7 (0.6–0.9)
Gastric cardia	19	1.0 (0.6–1.7)	132	1.0 (0.8–1.2)	40	2.7 (1.9–3.8)	519	4.5 (4.2–5.0)
Kidney	370	19.9 (17.9–22.1)	2,056	16.8 (16.0–17.5)	444	29.9 (27.0–33.0)	3,251	29.2 (28.2–30.3)
Liver	79	4.1 (3.3–5.2)	499	3.7 (3.4–4.1)	210	13.0 (11.2–15.1)	1,298	10.9 (10.3–11.6)
Multiple myeloma	289	15.2 (13.5–17.1)	760	5.7 (5.3–6.1)	253	18.4 (16.0–21.0)	945	8.4 (7.8–8.9)
Ovary*	206	11.1 (9.6–12.7)	1,681	13.8 (13.1–14.5)	-	-	-	-
Pancreas	399	22.3 (20.1–24.6)	1,662	12.2 (11.6–12.8)	295	21.3 (18.7–24.0)	1,887	16.6 (15.9–17.4)
Thyroid	256	13.4 (11.8–15.2)	2,273	22.4 (21.5–23.4)	62	4.1 (3.1–5.3)	767	7.4 (6.8–7.9)

* Female only.
† Excluded due to low counts.

Table 3. Estimated Population Attributable Fraction (PAF) (%) for Sex, Race, and Ethnicity by Cancer Type, Arkansas				
	Women		*Men*	
	Non-Hispanic Black, PAF%	Non-Hispanic White, PAF%	Non-Hispanic Black, PAF%	Non-Hispanic White, PAF%
Obesity-associated cancer type				
Breast, postmenopausal*	11.64%	9.68%	-	-
Colorectal	16.67%	13.99%	11.81%	14.63%
Corpus uteri*	60.53%	55.49%	-	-
Esophageal adenocarcinoma	†	52.72%	47.87%	54.02%
Gallbladder	25.78%	22.02%	18.87%	22.93%
Gastric cardia	7.22%	5.95%	4.95%	6.25%
Kidney	31.28%	27.01%	23.36%	28.05%
Liver	31.56%	27.27%	23.59%	28.32%
Multiple myeloma	12.11%	10.07%	8.44%	10.55%
Ovary*	17.74%	14.92%	-	-
Pancreas	21.97%	18.63%	15.86%	19.43%
Thyroid	5.12%	4.20%	3.48%	4.41%

* Female only
† Excluded due to low counts

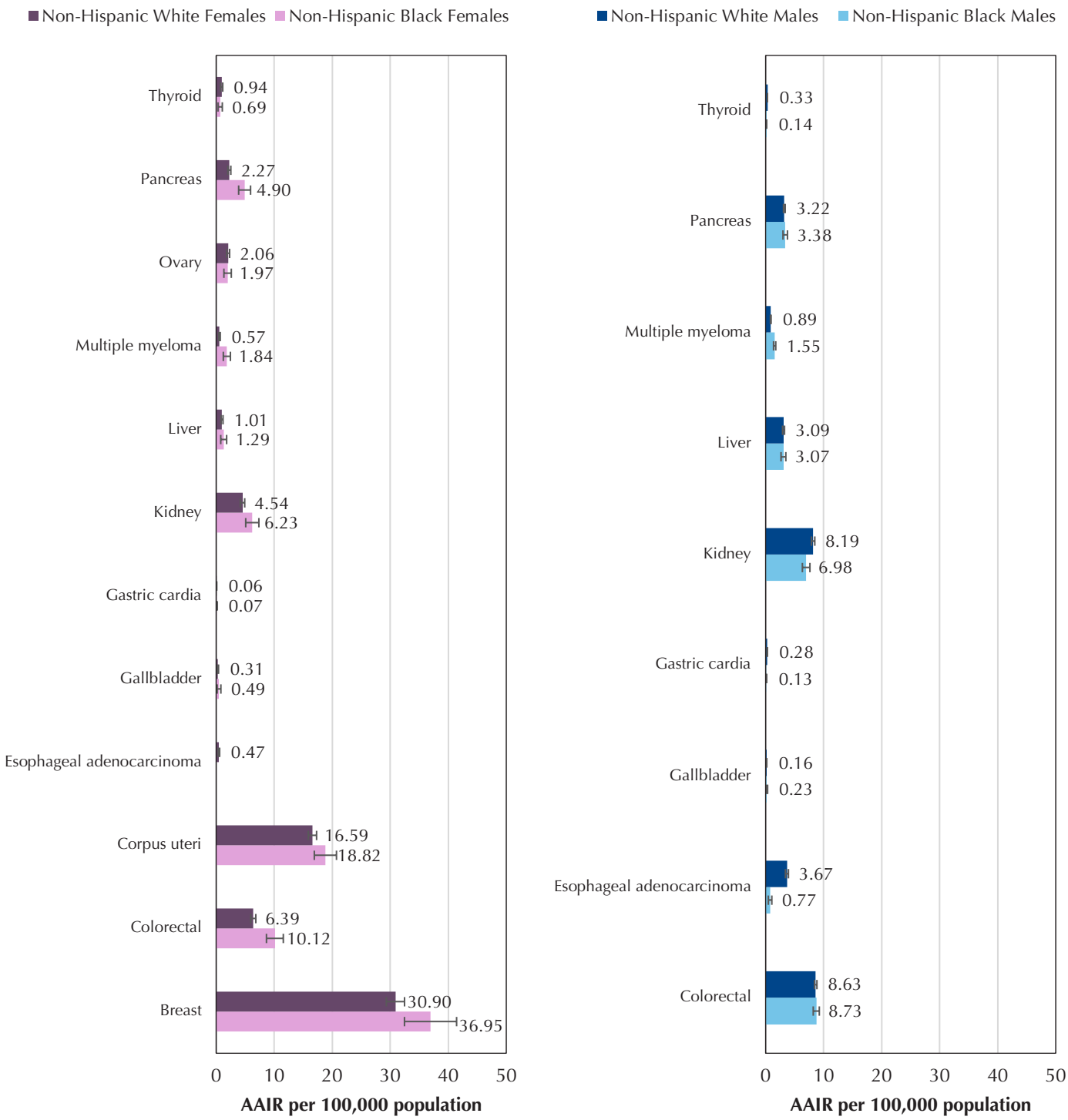
population.[8] Moreover, cancer rates in Arkansas closely mirror national patterns by race, ethnicity, and sex. For example, NH Black women had the highest obesity-associated cancer compared to all other major groups, though one group, NH White women, whose rate is higher in the United States overall than in Arkansas, showed slightly lower rates. Utilizing population-based cancer registry data made it possible to quantify Arkansas-specific obesity prevalence and to describe race and ethnicity-specific PAF by sex and cancer type. NH Black women are disproportionately impacted by obesity-attributable cancers, a disparity driven primarily by their higher incidence of breast and other female-specific cancers. Nevertheless, disparities by sex, race, and ethnicity were not statistically different across all types of cancer. Moreover, NH White men had a significantly higher rate of esophageal cancer. Greater use of smokeless tobacco products, especially among rural NH White men, may be one factor contributing to this disparity.[14]

This study has several implications for Arkansas. First, obesity is a well-established risk factor for both breast and colorectal cancer. Research shows that obesity increases the risk of breast cancer recurrence among patients and can negatively impact a survivor's quality of life.[15–17] Patients with obesity were more likely to be diagnosed with late-stage colorectal cancer, with a potential poor prognosis, and obesity has been linked to a 14% increase in colorectal-cancer-specific and all-cause mortality rates.[18] Timely screenings are recommended for breast and colorectal cancer in Arkansas. Although Arkansas ranks 31 in breast and 41 in colorectal cancer screening compared to other US states, evidence-based efforts are needed to increase screenings in the state, especially for at-risk populations with high rates of obesity.[19–21]

Second, current knowledge elucidates proposed biological mechanisms for an obesity-to-cancer progression.[1,22] For example, adipose tissue functions as an organ, releasing chemical mediators and enzymes, which can lead to excess production of estrogen. This excess production has been associated with a higher risk of developing female-specific cancers, including breast, endometrial, and ovarian cancer.[23–26] Another example is the increase of insulin and insulin-like growth factor 1 (IGF-1). High levels of insulin and IGF-1 are commonly seen among obese individuals and may contribute to the development of associated cancers, such as colon, renal, and endometrial cancer.[23,27,28] Finally, by weakening tumor immunity and altering the mechanical properties of the tissue surrounding growing tumors, obesity may increase cancer risk. Obesity has been linked with metastasis through various factors, such as adipokines, immune cell modulation, systemic inflammation, angiogenesis, metabolic changes, extracellular matrix alterations, and extracellular vesicles.[29,30] Considering Arkansas' statistically significant increase in adult obesity, there is a concerning association with obesity contributing to cancer-specific tumor development, growth, recurrence, and survival.[6,7,31–33] It is likely that the development of cancer in relation to obesity involves a variety of mechanisms; more research is needed to further understand them.

Third, weight management and interventions continue to be important not only in the prevention of cancer and reduction of overall mortality in cancer survivors, but also in comorbid illnesses and other chronic diseases.[6,34–36] Thus, it is important to address risk factors that lead to obesity.[37,38] The findings in this study suggest the continued need for obesity prevention initiatives at the state level that promote physical activity and access to preventive care. For example,

Figure 1. Estimated obesity-attributable cancers AAIR by sex, race and ethnicity, Arkansas, 2010-2019



it has been shown that, among type 2 diabetes patients, recently approved weight loss drugs (ie, glucagon-like peptide 1 receptor agonists [GLP-IRAs]) decrease the risk of certain obesity-associated cancers compared to insulin and metformin. However, lack of insurance coverage and high cost can hinder low-income, uninsured adults from utilizing GLP-IRAs.[39,40]

Moreover, evidence shows that combining structured exercise with dietary support for weight loss leads to greater weight loss than either exercise or diet alone.[41–45] This approach also has the most significant effect on blood biomarkers associated with common cancers, including insulin resistance, circulating levels of sex hormones, leptin, and inflammatory markers.[41–45] Considering that approximately 44.9% of Arkansans do not meet the recommended physical activity guideline and that 18.9% are food insecure, lack of access to safe and convenient places for physical activity and high-quality groceries may be contributing to obesity rates and disparities.[46–49] As not all populations have the same availability to resources for sustainable

weight loss or healthy weight maintenance, interventions may require a culturally tailored, multifaceted approach to improve access. Future work is needed to understand the variation in state- and demographic-specific cancers associated with obesity and to create focused approaches to decrease population obesity through regular physical activity and access to healthy foods in communities.

Limitations

Although the study has many strengths, it has at least three limitations. First, anthropometric data on weights and heights are self-reported in BRFSS, resulting in potential underreporting of obesity prevalence.[50] As a result, the burden of obesity-associated and attributable cancer in Arkansas is likely higher. Furthermore, BMI is commonly used as a surrogate measure for obesity but does not consider the varying distribution of excess adipose tissue in the body.[51] While there are studies using waist circumference and waist-hip ratio measures as better predictors of cancer risk, most research has utilized BMI due to its low cost and ease of measurement.[1] Second, the data assembled were not adequate to identify disparate age-adjusted incidence rates among all minoritized populations within the state, specifically excluding the state's growing Hispanic population and Marshallese community. Nor was data adequate to distinguish disparities across the urban-rural divide. Third, compared to Islami, et al., this study's findings showed noticeable discrepancies in PAF, which may be due to differences in the methodological approach and data sources.[5] Notably, colorectal, corpus uteri, esophageal adenocarcinoma, and ovarian cancer had a higher PAF, while the PAF was lower for gastric cardia and thyroid cancer. In general, states with lower cancer incidence rates that also have high rates of obesity have a higher PAF compared to other states, which may be the case for Arkansas.[5]

Conclusions

In conclusion, this study calculated a high obesity-attributable rate for cancers among women, especially among NH Black women for female-specific cancers. Men experienced a high obesity-attributable rate for kidney and liver cancer. Regular breast and colorectal screenings are suggested for individuals with obesity. Research also suggests that obesity is a contributing factor for cancer rates, which is a concern for Arkansas as a state that has consistently ranked among the highest in the nation in obesity prevalence. Addressing obesity in Arkansas through public health initiatives, such as access to healthy foods, access to physical activity opportunities, and education could play a crucial role in reducing the state's cancer burden.

Acknowledgements

We would like to acknowledge the contribution of Letitia de Graft-Johnson, DrPH, MHSA, LLB, Arkansas Survey Section Chief, Arkansas Department of Health, who provided Arkansas-specific obesity prevalence and expertise.

References

1. Pati S, Irfan W, Jameel A, Ahmed S, Shahid RK. Obesity and cancer: a current overview of epidemiology, pathogenesis, outcomes, and management. *Cancers*. 2023;15(2):485. doi:10.3390/cancers15020485
2. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body fatness and cancer—viewpoint of the IARC Working Group. *N Engl J Med*. 2016;375(8):794-798. doi:10.1056/NEJMsr1606602
3. Steele CB, Thomas CC, Henley SJ, et al. Vital signs: trends in incidence of cancers associated with overweight and obesity—United States, 2005–2014. *MMWR Morb Mortal Wkly Rep*. 2017;66(39):1052-1058. doi:10.15585/mmwr.mm6639e1
4. National Institutes of Health, National Cancer Institute. Common cancer types. Accessed May 10, 2024. **https://www.cancer.gov/types/common-cancers**
5. Islami F, Goding Sauer A, Gapstur SM, Jemal A. Proportion of cancer cases attributable to excess body weight by US state, 2011–2015. *JAMA Oncol*. 2019;5(3):384. doi:10.1001/jamaoncol.2018.5639
6. Aune D, Sen A, Prasad M, et al. BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *BMJ*. 2016;353:i2156. doi:10.1136/bmj.i2156
7. Petrelli F, Cortellini A, Indini A, et al. Association of obesity with survival outcomes in patients with cancer: a systematic review and meta-analysis. *JAMA Netw Open*. 2021;4(3):e213520. doi:10.1001/jamanetworkopen.2021.3520
8. National Program of Cancer Registries and Surveillance, Epidemiology and End Results Program SEER*Stat Database: NPCR and SEER Incidence - U.S. Cancer Statistics Public Use Research Database, 2022 Submission (2001–2020). Published online June 2023. Accessed at **https://www.cdc.gov/united-states-cancer-statistics/public-use**
9. Mansournia MA, Altman DG. Population attributable fraction. *BMJ*. 2018:360:k757. doi:10.1136/bmj.k757
10. Alexandra Lee, Michelle Cardel, William T Donahoo. Social and environmental factors influencing obesity. In: Kenneth R Feingold, Bradley Anawalt, Marc R Blackman, Alison Boyce, eds. Endotext [Internet]. Updated Oct. 12, 2019. MDText.com, Inc.; 2000. Accessed July 5, 2024. **https://www.ncbi.nlm.nih.gov/books/NBK278977/**
11. Clennin M, Reifler L, Goodman O, et al. Association of housing instability with obesity status among insured adults. *Am J Prev Med*. 2024 Sep;67(3):417-422. Published online April 2024:S0749379724001351. doi:10.1016/j.amepre.2024.04.008
12. Tiwari RC, Clegg LX, Zou Z. Efficient interval estimation for age-adjusted cancer rates. *Stat Methods Med Res*. 2006;15(6):547-569. doi:10.1177/0962280206070621
13. Division of Cancer Prevention and Control, Centers for Disease Control and Prevention. Definitions of risk factor-associated cancers. 2022. Accessed July 5, 2024. **https://www.cdc.gov/united-states-cancer-statistics/public-use/definitions-risk-factor-associated-cancers.html?CDC_AAref_Val=https://www.cdc.gov/cancer/uscs/public-use/predefined-seer-stat-variables.htm**
14. Cornelius ME, Loretan CG, Jamal A, et al. Tobacco product use among adults—United States, 2021. *MMWR Morb Mortal Wkly Rep*. 2023;72(18):475-483. doi:10.15585/mmwr.mm7218a1
15. Kang C, LeRoith D, Gallagher EJ. Diabetes, obesity, and breast cancer. *Endocrinology*. 2018;159(11):3801-3812. doi:10.1210/en.2018-00574
16. Mehta LS, Watson KE, Barac A, et al. Cardiovascular disease and breast cancer: where these entities intersect: a scientific statement from the American Heart Association. *Circulation*. 2018;137(8). doi:10.1161/CIR.0000000000000556
17. Connor AE, Baumgartner RN, Pinkston CM, Boone SD, Baumgartner KB. Obesity, ethnicity, and quality of life among breast cancer survivors and women without breast cancer: the long-term quality of life follow-up study. *Cancer Causes Control*. 2016;27(1):115-124. doi:10.1007/s10552-015-0688-z
18. Doleman B, Mills KT, Lim S, Zelhart MD, Gagliardi G. Body mass index and colorectal cancer prognosis: a systematic review and meta-analysis. *Tech Coloproctol*. 2016;20(8):517-535. doi:10.1007/s10151-016-1498-3

19. America's Health Rankings, United Health Foundation. *2024 Annual Report—Arkansas*. https://www.americashealthrankings.org/learn/reports/2024-annual-report/state-summaries-arkansas
20. State Cancer Profiles, National Cancer Institute. Screening and risk factors by state, had a mammogram in past 2 years. https://statecancerprofiles.cancer.gov/map/map.withimage.php?00&state&139&999&00&2&05&0&1&5&0#results
21. State Cancer Profiles, National Cancer Institute. Screening and risk factors by state, received at least one recommended CRC test. https://statecancerprofiles.cancer.gov/map/map.withimage.php?00&state&845&999&00&0&521&0&1&5&0#results
22. Avgerinos KI, Spyrou N, Mantzoros CS, Dalamaga M. Obesity and cancer risk: Emerging biological mechanisms and perspectives. *Metabolism*. 2019;92:121-135. doi:10.1016/j.metabol.2018.11.001
23. Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat Rev Cancer*. 2015;15(8):484-498. doi:10.1038/nrc3967
24. Bhardwaj P, Au CC, Benito-Martin A, et al. Estrogens and breast cancer: Mechanisms involved in obesity-related development, growth and progression. *J Steroid Biochem Mol Biol*. 2019;189:161-170. doi:10.1016/j.jsbmb.2019.03.002
25. Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer—mechanisms underlying tumour progression and recurrence. *Nat Rev Endocrinol*. 2014;10(8):455-465. doi:10.1038/nrendo.2014.94
26. Friedenreich CM, Ryder-Burbidge C, McNeil J. Physical activity, obesity and sedentary behavior in cancer etiology: epidemiologic evidence and biologic mechanisms. *Mol Oncol*. 2021;15(3):790-800. doi:10.1002/1878-0261.12772
27. Liu XZ, Pedersen L, Halberg N. Cellular mechanisms linking cancers to obesity. *Cell Stress*. 2021;5(5):55-72. doi:10.15698/cst2021.05.248
28. Shahid RK, Ahmed S, Le D, Yadav S. Diabetes and cancer: risk, challenges, management and outcomes. *Cancers*. 2021;13(22):5735. doi:10.3390/cancers13225735
29. Seo BR, Bhardwaj P, Choi S, et al. Obesity-dependent changes in interstitial ECM mechanics promote breast tumorigenesis. *Sci Transl Med*. 2015;7(301). doi:10.1126/scitranslmed.3010467
30. Annett S, Moore G, Robson T. Obesity and cancer metastasis: molecular and translational perspectives. *Cancers*. 2020;12(12):3798. doi:10.3390/cancers12123798
31. Ligibel JA, Alfano CM, Hershman D, et al. Recommendations for obesity clinical trials in cancer survivors: American Society of Clinical Oncology statement. *J Clin Oncol*. 2015;33(33):3961-3967. doi:10.1200/JCO.2015.63.1440
32. Lazarus E, Bays HE. Cancer and obesity: an Obesity Medicine Association (OMA) clinical practice statement (CPS) 2022. *Obes Pillars*. 2022;3:100026. doi:10.1016/j.obpill.2022.100026
33. Slawinski CGV, Barriuso J, Guo H, Renehan AG. Obesity and cancer treatment outcomes: interpreting the complex evidence. *Clin Oncol*. 2020;32(9):591-608. doi:10.1016/j.clon.2020.05.004
34. Grundy SM, Williams C, Vega GL. Upper body fat predicts metabolic syndrome similarly in men and women. *Eur J Clin Investigation*. 2018;48(7):e12941. doi:10.1111/eci.12941
35. Alamuddin N, Bakizada Z, Wadden TA. Management of obesity. *J Clin Oncol*. 2016;34(35):4295-4305. doi:10.1200/JCO.2016.66.8806
36. Ligibel JA, Bohlke K, May AM, et al. Exercise, diet, and weight management during cancer treatment: ASCO guideline. *J Clin Oncol*. 2022;40(22):2491-2507. doi:10.1200/JCO.22.00687
37. Rock CL, Thomson C, Gansler T, et al. American Cancer Society guideline for diet and physical activity for cancer prevention. *CA Cancer J Clin*. 2020;70(4):245-271. doi:10.3322/caac.21591
38. Rock CL, Thomson CA, Sullivan KR, et al. American Cancer Society nutrition and physical activity guideline for cancer survivors. *CA Cancer J Clin*. 2022;72(3):230-262. doi:10.3322/caac.21719
39. Wang L, Xu R, Kaelber DC, Berger NA. Glucagon-like peptide 1 receptor agonists and 13 obesity-associated cancers in patients with type 2 diabetes. *JAMA Netw Open*. 2024;7(7):e2421305. doi:10.1001/jamanetworkopen.2024.21305
40. Vokinger KN, Nussli E, Dusetzina SB. Access to GLP-1 weight loss drugs in the US, Canada, Switzerland, and Germany. *JAMA Intern Med*. 2024;184(9):1002. doi:10.1001/jamainternmed.2024.2559
41. Van Gemert WAm, Schuit AJ, Van Der Palen J, et al. Effect of weight loss, with or without exercise, on body composition and sex hormones in postmenopausal women: the SHAPE-2 trial. *Breast Cancer Res*. 2015;17(1):120. doi:10.1186/s13058-015-0633-9
42. Campbell KL, Foster-Schubert KE, Alfano CM, et al. Reduced-calorie dietary weight loss, exercise, and sex hormones in postmenopausal women: randomized controlled trial. *J Clin Oncol*. 2012;30(19):2314-2326. doi:10.1200/JCO.2011.37.9792
43. Foster-Schubert KE, Alfano CM, Duggan CR, et al. Effect of diet and exercise, alone or combined, on weight and body composition in overweight-to-obese postmenopausal women. *Obesity*. 2012;20(8):1628-1638. doi:10.1038/oby.2011.76
44. Imayama I, Ulrich CM, Alfano CM, et al. Effects of a caloric restriction weight loss diet and exercise on inflammatory biomarkers in overweight/obese postmenopausal women: a randomized controlled trial. *Cancer Res*. 2012;72(9):2314-2326. doi:10.1158/0008-5472.CAN-11-3092
45. Mason C, Foster-Schubert KE, Imayama I, et al. Dietary weight loss and exercise effects on insulin resistance in postmenopausal women. *Am J Prev Med*. 2011;41(4):366-375. doi:10.1016/j.amepre.2011.06.042
46. Carly Mallenbaum. Study: Most Arkansans aren't getting enough exercise. Axios. March 14, 2024. https://www.axios.com/local/nw-arkansas/2024/03/14/arkansas-exercise-enough-map-health
47. Matthew P. Rabbitt, Madeline Reed-Jones, Laura J. Hales, Michael P. Burke. *Household Food Security in the United States in 2023*. U.S. Department of Agriculture, Economic Research Service; 2024. https://arkansasadvocate.com/wp-content/uploads/2024/09/Household-Food-Insecurity-in-the-United-States-in-2023.pdf
48. Myers CA, Mire EF, Katzmarzyk PT. Trends in Adiposity and Food Insecurity Among US Adults. *JAMA Netw Open*. 2020;3(8):e2012767. doi:10.1001/jamanetworkopen.2020.12767
49. Baillot A, Chenail S, Barros Polita N, et al. Physical activity motives, barriers, and preferences in people with obesity: A systematic review. Kumar S, ed. *PLoS ONE*. 2021;16(6):e0253114. doi:10.1371/journal.pone.0253114
50. Olfert MD, Barr ML, Charlier CM, et al. Self-reported vs. measured height, weight, and BMI in young adults. *Int J Environ Res Public Health*. 2018;15(10):2216. doi:10.3390/ijerph15102216
51. Lim S, Meigs JB. Ectopic fat and cardiometabolic and vascular risk. *Int J Cardiol*. 2013;169(3):166-176. doi:10.1016/j.ijcard.2013.08.077

Original Article

Inferring Unknown Race in Central Cancer Registries

Francis P. Boscoe, Ph.D.[a]

Abstract: Completeness of race information is a criterion for data certification among United States central cancer registries. This paper presents a method for reducing unknown race information by as much as 75%, with 95% accuracy, race-specific sensitivity of 81–99%, and race-specific positive predictive value of 88–97%. The method, Bayesian Improved Surname Geocoding (BISG), has been in wide use in the social sciences and public health for more than 15 years. We use the publicly available North Carolina voter rolls as a proxy for cancer patients, drawing a sample of these persons that mimics the national distribution of cancer incidence by race and ethnicity. BISG has the potential to increase the accuracy of race-specific cancer incidence rates by as much as 3% in registries with the highest levels of missingness; in other registries, the effects will be negligible. The method has been incorporated into freely available computer code.

Key words: *Bayesian Improved Surname Geocoding, ethnicity, missingness, race, race-specific rates*

Introduction

Patient race is a core component of cancer surveillance.[1–5] In central cancer registries in the United States, for cancer cases diagnosed from 2017–2021, race is missing for about 1% of patients.[6] California has the highest rate of missingness at 3.4%. Four other Western states (Arizona, New Mexico, Washington, and Oregon) have rates of 1.7–1.8%.[6] For the registries which comprise California, Greater Bay has 2.0% with missing race, Los Angeles 3.2%, and Greater California 3.9%.[6] The national gold standard for missing race is 3% and the silver standard is 5%, meaning that most of California is only at the silver level for these years.[7] This paper describes a method for assigning race based on patients' full names and residential locations that has the potential to reassign approximately 3/4 of those with missing race, with over 95% accuracy and race-specific sensitivity and with positive predictive value of 80–99%, which would have the effect of raising all of California to the gold level for this measure.

There is a long tradition within public health and the social sciences of using available demographic information to infer missing information. Spanish surname lists have existed since the 1950s.[8] Other types of name lists appeared around the turn of the millennium as information technology made such lists feasible. Lauderdale and Kestenbaum, for example, developed a surname list for 6 Asian ethnic groups using Social Security Administration records.[9] That list was later expanded to include first names, augmented by Medicare records.[10] Wong et al. applied self-reported race and ethnicity information from a large health-care provider to categorize names into 14 exclusive race and Asian ethnicity categories and Hispanic ethnicity.[11–12] Nasseri was the first to develop a Middle Eastern surname list, also derived from Social Security Administration records.[13] Nearly all central cancer registries in the United States have been using birthplaces and surnames to infer Hispanic ethnicity and specific Asian and Pacific Islander race since the early 2000s.

The United States Census Bureau's release of the first population-based surname list following the 2000 census represented a major advance.[14] This list contains all surnames occurring at least 100 times in the nation, incorporating over 90% of the population, with each name classified by the number in 6 race/ethnic groups: White, Black, Asian or Pacific Islander (API), American Indian/Alaska Native (AIAN), Hispanic, and multiple races, with the percentages of these 6 groups summing to 100. Another edition of this file was released following the 2010 census.[15] More comprehensive first name lists have become available more recently, including a list developed by Tzioumis from mortgage application databases[16] and a much larger list by Rosenman et al. from statewide voter rolls,[17] the latter of which was used in the current analysis.

Supplementing name lists with patient address information yields substantially more complete and accurate inferences about race and ethnicity. The United States is strongly racially segregated; all else being equal, a person with the surname Lee living in a heavily Asian neighborhood is more likely to be Asian than a person with this surname in a rural area. The most widely implemented version of this approach is known as Bayesian Improved Surname Geocoding (BISG);[18–24] which a recent review described as "ubiquitous."[25] Bayesian refers to applying Bayes' rule to find the conditional probability that an individual belongs to a given race/ethnicity given their name, residential location, and potentially other demographic variables.[26] The method has been incorporated into a freely available R software package called wru, which uses available name lists and census data on the racial and ethnic composition of small geographic areas to make individual predictions about race and ethnicity.[27]

---

[a]Francis P. Boscoe, Ph.D., Pumphandle, LLC, Camden, Maine

Address correspondence to Francis P. Boscoe, Ph.D., Pumphandle, LLC, 6 Wilson Ave., Camden, Maine, 04843.
Email: francis.boscoe@pumphandle.consulting

Francis P. Boscoe completed this analysis while serving as a consultant for NAACCR under the supervision of Recinda Sherman.

Methods and Materials

We assessed 3 methods for imputing missing race information for cancer patients. The first and simplest used only the population-based list of surnames occurring at least 100 times in the 2010 United States census. We calculated percentages for people of non-Hispanic White, Black, API, and other racial backgrounds, where the category of other was the sum of AIAN and multiple races. Non-Hispanic percentages were necessary to match the format of cancer registry data, where race and Hispanic ethnicity are listed separately. To illustrate, the name Martin has a distribution that is 75% White, 16% Black, 1% API, 1% AIAN, 2% multiple races, and 6% Hispanic, rounded to the nearest whole percent. After removing Hispanic persons, the distribution became 79% White, 17% Black, 1% API, 1% AIAN, and 2% multiple races.

We then tallied which names were strongly associated with White, Black or API race, using a threshold of either 75% or 85%. The surname Martin, for example, meets the 75% threshold for White but not the 85% threshold. Specifically, this is a positive predictive value threshold.

The second method was like the first, but augmented with first, middle, and last names from the voter rolls from 6 states (Alabama, Florida, Georgia, Louisiana, North Carolina, and South Carolina).[23–24] To our knowledge, this is the largest first- and middle-name list in existence. The threshold for inclusion was 25 occurrences across 7 vintages of the file from each state. Because a person could be counted as many as 7 times, the effective threshold for inclusion could be as low as 4 persons. Though this file uses only 6 states representing 17.2% of the population — further reduced because it only included those registered to vote — it contains roughly double the number of names that are on the census list. Between the first, middle, and last names, we selected the name with the highest association with either White, Black or API race, and tested these against the 75% and 85% thresholds as in the first method.

For the third method, we used the predict_race() function in the WRU package, using both the 2010 census and voter-roll name lists from the first 2 methods (both of which are incorporated into this package), at the geographic level of county. Again, the results were filtered by the 75% and 85% thresholds.

Simulated Patient Sample

As we had no access to the names of cancer patients, we instead tested our methods using a simulated sample of 500,000 names with known race and ethnicity contained in the publicly-available North Carolina voter rolls.[28] To simulate the patient sample, we first restricted the North Carolina file to those with recorded race and ethnicity among the races listed. The raw file contains first, middle, and last names, residential address, age, sex, political party, and other demographic information for 8,789,530 registered voters, but only 5,713,228 records included self-reported race and ethnicity. We used the following race categories: White, Black, Asian, Pacific Islander, American Indian, multiple races, other race, and unknown race. Ethnic categories are Hispanic, non-Hispanic, and unknown. Thus, we limited the analysis to those whose records indicated White, Black, Asian or Pacific Islander race, and who also indicated whether or not they were Hispanic.

American Indian persons were excluded from the analysis because of small numbers. While the Eastern Band of Cherokee is a federally-recognized tribe in North Carolina, and several other tribes and bands are recognized by the state, collectively they account for less than 1% of the voter roll. In addition, inspection of those with a race code of American Indian in the North Carolina data revealed that many had typically Asian Indian names.

The sample was weighted to match the racial and ethnic composition of White, Black, and API cancer patients in the United States as a whole for diagnosis years 2017–2021. More specifically, the weighting included all states plus the District of Columbia, but excluded Kansas, for which county-level cancer incidence data were not available. All US territories were also excluded. A comparison between the North Carolina data and the sample is given in Table 1. The sample has a similar proportion of non-Hispanic White persons, about half as many Black persons, more than double the share of API persons, and over 6 times the number of Hispanic persons contained in the North Carolina data.

The sample was further matched to the racial and ethnic composition of each county. For example, Cook County (Chicago), Illinois reported 33,914 cancer cases among Black non-Hispanic persons, which is about 0.4% of the total cancers in the US for all races/ethnicities. Thus,

Table 1. Comparison of North Carolina voter data with the simulated patient sample

Race/ethnicity	*North Carolina*	*Simulated Patient Sample*
White, non-Hispanic	4,313,063 (75.5%)	382,924 (76.6%)
Black, non-Hispanic	1,227,659 (21.5%)	56,497 (11.3%)
Asian or Pacific Islander, non-Hispanic	94,545 (1.7%)	18,246 (3.6%)
White, Hispanic	65,220 (1.1%)	40,540 (8.1%)
Black, Hispanic	11,385 (0.2%)	1,424 (0.3%)
Asian or Pacific Islander, Hispanic	1,356 (0.02%)	369 (0.1%)
Total	5,713,228 (100.0%)	500,000 (100.0%)

Note: The North Carolina file was restricted to those with known race and ethnicity among the races listed. The raw file contains 8,789,530 records.

Table 2. Comparison of three methods for assigning race using name lists and county of residence

Method	*Threshold*	*% Recoded*	*Accuracy*	*Sensitivity—White (%)*	*Sensitivity—Black (%)*	*Sensitivity—API (%)*	*Specificity—White (%)*	*Specificity—Black (%)*
Census surnames only	0.75	45.1	92.5	100	9	80	91	93
Census surnames only	0.85	26.0	96.0	100	13	87	94	97
Census surnames + Voter roll names	0.75	92.8	92.1	99	39	68	91	93
Census surnames + Voter roll names	0.85	73.0	95.0	99	54	73	93	96
wru package	0.75	79.9	96.0	98	75	93	87	98
wru package	0.85	71.1	97.4	99	80	94	92	98

Table 2, cont. Comparison of three methods for assigning race using name lists and county of residence

Method	*Specificity—API (%)*	*PPV—White (%)*	*PPV—Black (%)*	*PPV—API (%)*	*Cohen's Kappa—White*	*Cohen's Kappa—Black*	*Cohen's Kappa—API*
Census surnames only	99.0	93	80	86	0.46	0.15	0.86
Census surnames only	99.0	95	88	88	0.70	0.22	0.90
Census surnames + Voter roll names	99.0	92	87	96	0.57	0.51	0.79
Census surnames + Voter roll names	99.0	95	90	96	0.70	0.65	0.82
wru package	99.7	97	88	82	0.83	0.79	0.86
wru package	99.8	98	92	87	0.88	0.85	0.90

0.4% of the sample (about 2,000) consisted of Black non-Hispanic persons assigned to Cook County, Illinois.

Evaluation

For each of the 3 approaches just described, we compared the known race with the predicted race, and tabulated the sensitivity, specificity, and positive predictive value for White, Black, and API race, along with the overall accuracy and overall percentage that were assigned to a more specific race. Sensitivity in this case refers to the proportion of a particular race who were correctly assigned that race. Specificity refers to the proportion who do not belong to a racial group and were correctly not assigned to that group. Positive predictive value refers to the proportion who were correctly assigned a race. Accuracy is the ratio of those assigned a correct race to those assigned either a correct or incorrect race.

Results

For each of the 3 methods, Table 2 gives the sensitivity, specificity, positive predictive value, accuracy, percent recoded, and Cohen's kappa statistic. For the first 2 methods, the sensitivity for Black race was poor, with most Black persons miscoded as White. The third approach, which additionally incorporated county of residence, yielded sensitivity of at least 75% and positive predictive value of at least 80% for all racial groups, along with more than 70% recoded and an overall accuracy of at least 95%. Further stratifying the WRU package results by Hispanic status revealed that the high-quality results were confined to non-Hispanic persons (Table 3). For Hispanic persons, relatively few were assigned a racial code, and with poorer accuracy. This is also reflected in the poor specificity for White Hispanic persons. Cohen's kappa closely tracked accuracy in both tables.

In order not to be limited to only 2 name-based thresholds, we calculated sensitivity, positive predictive value, accuracy, and percent recoded for all whole-number thresholds between 50% and 99%, limiting the analysis to non-Hispanic ethnicity. These results are presented graphically in Figures 1 and 2. Figure 1 shows that the results are best for White race, followed by API and Black race. Sensitivity for Black race is the lagging measure across nearly the entire range of threshold values.

Los Angeles County would be among the locations most impacted from this method, so we calculated this impact, assuming that cancer patients with unknown race in Los Angeles County followed the distribution of those with known race and that Los Angeles County otherwise mimicked the national results. The 1,578 Hispanic patients

Table 3. Results from WRU Package, Stratified by Hispanic Ethnicity								
Method	*Threshold*	*% Recoded*	*Accuracy*	*Sensitivity—White (%)*	*Sensitivity—Black (%)*	*Sensitivity—API (%)*	*Specificity—White (%)*	*Specificity—Black (%)*
Non-Hispanic	0.75	83.8	96.6	99	76	93	92	98
Non-Hispanic	0.85	75.0	97.8	99	81	95	95	98
Hispanic	0.75	37.4	82.9	84	66	93	19	98
Hispanic	0.85	28.2	86.2	87	70	94	23	99

Table 3, cont. Results from WRU Package, Stratified by Hispanic Ethnicity							
Method	*Specificity—API (%)*	*PPV—White (%)*	*PPV—Black (%)*	*PPV—API (%)*	*Cohen's Kappa—White*	*Cohen's Kappa—Black*	*Cohen's Kappa—API*
Non-Hispanic	99.7	97	90	91	0.85	0.80	0.92
Non-Hispanic	99.8	98	94	94	0.90	0.86	0.98
Hispanic	99.5	99	38	6	0.27	0.45	0.10
Hispanic	99.6	99	45	9	0.33	0.52	0.14

Figure 1. Sensitivity and Positive Predictive Value by Race in Simulated National Sample, by Positive Predictive Value Threshold

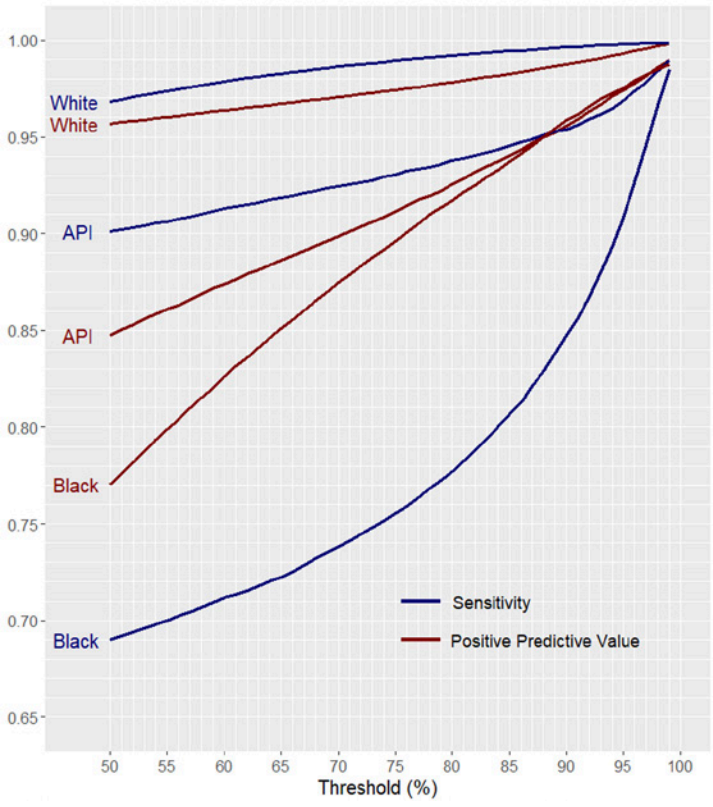
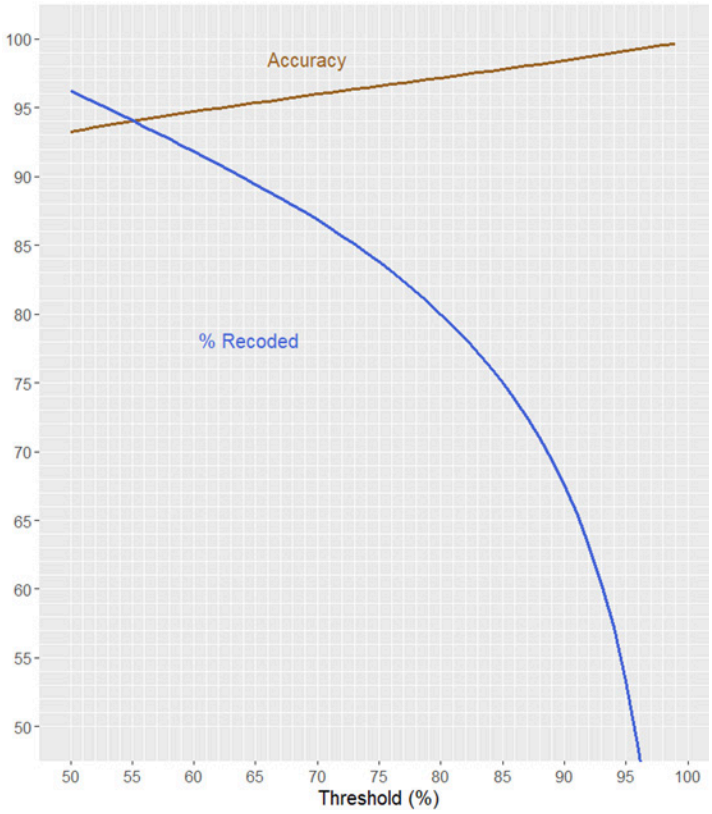


Figure 2. Overall Accuracy and Percent Recoded in Simulated National Sample, by Positive Predictive Value Threshold



with unknown race would remain as such, while 75% of the 5,008 non-Hispanic patients would be assigned a known race, leaving a total of 2,830 with unknown race, or 1.4%, well below the 3% threshold for gold certification. Crude rates for White cancer patients would be expected to increase by 1.8%, for Black patients by 1.4%, and for Asian patients by 2.7%.

Discussion

We have shown that the predict_race() function within the WRU package in R gives good results for predicting whether a person identifies as White, Black, or API with knowledge of their full name and county of residence. The simpler approaches of using only name lists (either the census surname list alone or in combination with first,

middle, and surname name lists developed from voter rolls) do not yield acceptable results for predicting Black race, among other weaknesses. While the name-list-only methods here function as straw men, in the sense that the literature review has already established the superiority of BISG, we included them here because they reflect current practice in cancer registration, which has long used name-based algorithms to ascertain Hispanic ethnicity (NAACCR Hispanic Identification Algorithm, or NHIA) and specific API ethnicity conditional on being API (NAACCR Asian and Pacific Islander Identification Algorithm, or NAPIIA).[29–30] This is not to malign these other methods: The reason name-only methods are good at identifying Hispanic and API individuals, but not Black individuals, is because Hispanic and API names have much less overlap with non-Hispanic White names than Black names do. Even so, our study shows that a BISG approach can improve the results for all groups, not just for those of Black race.

If central cancer registries were to adopt this method, a question remains as to which threshold should be chosen. If the goal is simply to minimize the number with unknown race, then the threshold should be low. However, registries value both completeness and accuracy. For NHIA and NAPIIA, a threshold of 75% has been in place since their inception. At this level, in the current study, the weakest result is a sensitivity of 0.76 for Black race, meaning that 24% of Black persons are misclassified, mostly as White. Indeed, other researchers have identified this specific issue as the largest weakness of BISG.[20] We propose the 85% threshold as a more conservative alternative. At this threshold, the sensitivity for Black race improves to 0.81, along with improvements in all other measures, at the cost of reclassifying 75% of non-Hispanic persons, down from 84%.

While this paper is concerned solely with cancer patients with a race code of 99, a BISG approach could also be used to enhance the existing NHIA and NAPIIA algorithms. For example, consider the surname Angel, which is 44% non-Hispanic White and 51% Hispanic according to the 2010 census, making it a "generally" Hispanic surname, and so reclassified as non-Hispanic.[30] In contrast, the predict_race() function in WRU gives a 99% probability that a person with this surname is Hispanic if they live in Maverick County, Texas, which is 95% Hispanic overall. Location thus matters.

There are several ways in which the method described here could be further improved. Subcounty-level geocoding would be superior to county-level, particularly in large urban counties with significant racial and ethnic segregation. For example, consider the name Jerold Armstrong, with both first and last names about 20% Black. The predict_race() function assigns this person a 51% probability of being Black if he lives in Philadelphia, Pennsylvania. In the Nicetown neighborhood within Philadelphia, however, the probability jumps to 95%; in the Chestnut Hill neighborhood, it drops to 13%. Including age and sex would also be expected to net small improvements.[26] While the documentation of the WRU package says these can be used as inputs, in the versions of the WRU package available at the time of this analysis (versions 3.03 and 3.04), this feature was not functional. The method could, in principle, be expanded to identify American Indian and Alaska Native persons; however, only 259 surnames are identified in the 2010 census as occurring in this group at least 85% of the time, only 6 of which occurred more than 1,000 times. In addition, since the method only used names from North Carolina, validation using actual cancer registry names is warranted. This would require as many registries as possible testing the method with their own data where race and ethnicity are already known and comparing the algorithmic results with the known results. To the degree that names in North Carolina are not a representative sample of names nationally, we anticipate that the registry data would outperform the North Carolina data, making the results we report here conservative. In addition, cancer registries often have information on the birth names of women who change their names when married. These names are absent from the voter rolls, providing another reason why these results are likely conservative.

Several authors have described bias inherent in the method and how it might be reduced.[25,31–32] This bias is evident in Table 3 and Figure 1, insofar as different race/ethnic groups have different levels of reclassification and accuracy. Moreover, people tend to cluster in families that share surnames and families tend to live nearer to one another than unrelated people, which violates the assumption of independence among variables. Methods for reducing this bias include machine learning classification approaches, the incorporation of additional covariates such as census-tract level income and home value, and a ranking-based approach to populate a 3-dimensional contingency table with race/ethnicity, geographic location, and surname as the 3 dimensions. Any of these enhancements would incur programming effort, whereas the WRU package is already functional, giving it a practical advantage.

In conclusion, these results support the adoption of a method for inferring race for cancer patients missing this information, one that is informed by full patient names and geographic locations. It would likely have the effect of moving registries not meeting the gold-certification standard for missing race comfortably below the threshold. Further testing using actual registry data is recommended before implementation.

References

1. Ellis L, Canchola AJ, Spiegel D, Ladabaum U, Haile R, Gomez SL. Racial and ethnic disparities in cancer survival: the contribution of tumor, sociodemographic, institutional, and neighborhood characteristics. *J Clin Oncol* 2018; 36(1):25-33.
2. Zavala VA, Bracci PM, Carethers JM, et al. Cancer health disparities in racial/ethnic minorities in the United States. *Br J Cancer* 2021; 124(2): 315-332.
3. Howlader N, Bhattacharya M, Scoppa S, et al. Cancer and COVID-19: US cancer incidence rates during the first year of the pandemic. *J Natl Cancer Inst* 2024;116(2): 208-215.
4. Vince RA Jr, Eyrich NW, Mahal BA, Stensland K, Schaeffer EM, Spratt DE. Reporting of racial health disparities research: are we making progress? *J Clin Oncol* 2022; 40(1): 8-11.
5. Martinez RAM, Andrabi N, Goodwin AN, Wilbur RE, Smith NR, Zivich PN. Conceptualization, operationalization, and utilization of race and ethnicity in major epidemiology ojurnals, 1995-2018: a systematic review. *Am J Epidemiol* 2023; 192(3): 483-496.

6. SEER*Stat Database: NAACCR Incidence Data - CFD 2023, 1995-2021, for U.S. and CDN, Boscoe - NHAPIIA TF (which includes data from CDC's National Program of Cancer Registries (NPCR), CCR's Provincial and Territorial Registries, and the NCI's Surveillance, Epidemiology and End Results (SEER) Registries), certified by the North American Association of Central Cancer Registries (NAACCR) as meeting high-quality incidence data standards for the specified time periods, submitted December 2023.
7. North American Association of Central Cancer Registries. Certification levels & recognition logos. Accessed March 5, 2025. **https://www.naaccr.org/certification-criteria/**
8. Perkins RC. Evaluating the Passel-Word Spanish surname list: 1990 decennial census post-enumeration survey results. United States Bureau of the Census, Population Division, Technical Working Paper Series No. 4.
9. Lauderdale DS, Kestenbaum B. Asian American ethnic identification by surname. *Popul Res Policy Rev* 2000; 19: 283-300.
10. Lauderdale DS, Kestenbaum B. Mortality rates of elderly Asian American populations based on Medicare and Social Security data. *Demography* 2002; 39(3): 529-40.
11. Palaniappan LP, Wong EC, Shin JJ, Moreno MR, Otero-Sabogal R. Collecting patient race/ethnicity and primary language data in ambulatory care settings: a case study in methodology. *Health Serv Res* 2009; 44(5 Pt 1): 1750-1761.
12. Wong EC, Palaniappan LP, Lauderdale DS. Using name lists to infer Asian racial/ethnic subgroups in the healthcare setting. *Med Care* 2010; 48(6): 540-546.
13. Nasseri K. Construction and validation of a list of common Middle Eastern surnames for epidemiological research. *Cancer Detect Prev* 2007; 31(5): 424-429.
14. Word DL, Coleman CD, Nunziata R, Kominski R. Demographic aspects of surnames from Census 2000. Accessed March 10, 2025. **https://www2.census.gov/topics/genealogy/2000surnames/surnames.pdf**
15. Comenetz J. 2016. Frequently occurring surnames in the 2010 Census. Accessed March 10, 2025. **https://www2.census.gov/topics/genealogy/2010surnames/surnames.pdf**
16. Tzioumis K. Demographic aspects of first names. *Sci Data* 2018; 5: 1-9.
17. Rosenman ETR, Olivella S, Imai K. Race and ethnicity data for first, middle, and surnames. *Sci Data* 2023; 10(1): 299.
18. Fremont AM, Bierman A, Wickstrom SL, et al. Use of geocoding in managed care settings to identify quality disparities. *Health Aff* (Millwood) 2005; 24(2): 516-26.
19. Fiscella K, Fremont AM. Use of geocoding and surname analysis to estimate race and ethnicity. *Health Serv Res* 2006; 41: 1482-1500.
20. Elliott MN, Fremont A, Morrison PA, Pantoja P, Lurie N. A new method for estimating race/ethnicity and associated disparities where administrative records lack self-reported race/ethnicity. *Health Serv Res* 2008; 43(5 Pt 1): 1722-36.
21. Elliott MN, Morrison PA, Fremont A, McCaffrey DF, Pantoja P, Lurie N. Using the Census Bureau's surname list to improve estimates of race/ethnicity and associated disparities. *Health Serv Outcomes Res Method* 2009; 9: 69-83.
22. Voicu I. Using first name information to improve race and ethnicity classification. *Stat Public Policy* 2018; 5(1): 1-13.
23. Deluca K, Curiel JA. Validating the applicability of Bayesian inference with surname and geocoding to Congressional Redistricting. *Political Anal* 2022; 31(3): 465-471.
24. Imai K, Olivella S, Rosenman ETR. Addressing Census data problems in race imputation via fully Bayesian Improved Surname Geocoding and name supplements. *Sci Adv* 2022; 8(49): eadc9824.
25. Greengard P, Gelman A. A calibrated BISG for inferring race from surname and geolocation. *J R Stat Soc* Series A: Statistics in Society 2025; in press.
26. Imai K, Khanna K. Improving ecologic inference by predicting individual ethnicity from voter registration records. *Political Anal* 2016; 24: 263-272.
27. Khanna K, Bertelsen B, Olivella S, Rosenman E, Hayes AR, Imai K. 2024. WRU: Who are you? Bayesian prediction of racial category using surname and geolocation. R package version 3.03. **https://cloud.r-project.org/web/packages/wru/index.html**
28. North Carolina State Board of Elections. Statewide Voter Registration [data file]. Accessed March 10, 2025. **https://www.ncsbe.gov/results-data/voter-registration-data**
29. NAACCR Race and Ethnicity Work Group. 2010. *NAACCR Asian Pacific Islander Identification Algorithm* [NAPIIA v1.2.1]. Springfield, IL: North American Association of Central Cancer Registries.
30. NAACCR Race and Ethnicity Work Group. 2011. *NAACCR Guideline for Enhancing Hispanic/Latino Identification: Revised NAACCR Hispanic/Latino Identification Algorithm* [NHIA v2.2.1]. Springfield, IL: North American Association of Central Cancer Registries.
31. Decter-Frain A. How should we proxy for race/ethnicity? Comparing Bayesian Improved Surname Geocoding to Machine Learning Methods. *arXiv* 2022; 2206.14583v2.
32. Argyle L, Barber M. Misclassification and bias in predictions of individual ethnicity from administrative records. *Am Political Sci Rev* 2024; 118(2): 1058-1066.

Elekta

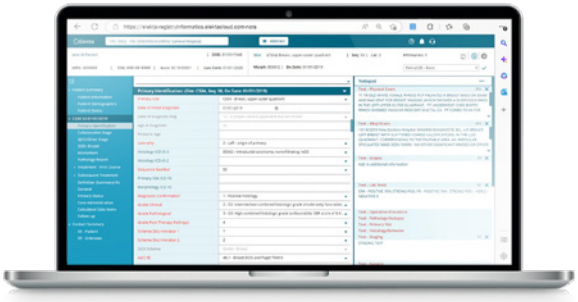
Elekta ONE® Registry Informatics

Drive your cancer registry workflow forward

"When I run a report in Registry Informatics, it is really fast. Everything I want is right there. Things I could not do with METRIQ, I am able to do with Registry Informatics. It is very easy."

Hope Valcarcel, BS, RHIT, ODS
Oncology Data Management Consultant

Experience the power



elekta.com/registryinformatics

Coalition Collaboration to Improve Timely Cancer Registry Data Reporting

Dana Doyle, MPH

Summary

The Arizona Cancer Registry (ACR) is collaborating with the Arizona Cancer Coalition to address delinquent case reporting by mandated reporting facilities. In 1988, the Arizona Revised Statute §36- 133 was amended to mandate reporting of cancer cases to the ACR and these rules were put into effect on January 1, 1992. Per the revised statute, hospital facilities are required to report a case to the ACR within 180 calendar days of the date of first release. Physicians and clinics are required to report a case within 30 to 90 calendar days depending on the clinic size.

Between 2023 and 2024, only 27% of cases received by the ACR were reported on time according to regulation. Hospital systems had the most noncompliant case reporting. In response, the ACR collaborated with the Arizona Cancer Coalition to create an objective and strategies to improve timely cancer reporting in the new 2024 to 2029 Arizona Cancer Plan.

Challenges

- The Arizona Revised Statute does not provide any penalties for delinquent reporting. The options for the ACR to address the issue of delinquent case reporting are to work directly with facilities and/or the Arizona Department of Health Services administrative counsel and licensing department.
- Reporting facilities need help to identify barriers to reporting cancer cases on time to the ACR. The reporting facilities can begin to address these barriers internally and in collaboration with the ACR with the understanding that solutions such as training and staffing may be costly and time intensive.
- If reporting facilities are delinquent in case reporting, it degrades the overall completeness and accuracy of ACR data.

Solution

The ACR and members of the Arizona Cancer Coalition Policy work group created an objective to increase the percentage of cancer cases reported on time and in accordance with regulation from 27% to 75%. The strategies to address the objective include:

- Developing materials on the importance of cancer reporting and disseminating them to hospital administrators.
- Developing and implementing a quality improvement process that shares hospital cancer reporting information.
- Increasing electronic cancer reporting from health care professionals and clinics.

The goal of working with the coalition is to raise awareness and garner support that encourages reporting facilities to work internally, and with the ACR, to address the barriers to on-time cancer reporting.

Results

- The results of this work are ongoing as the registry engages with coalition members to implement the strategies in the Arizona Cancer Plan. These strategies aim to encourage hospitals to improve on-time case reporting and promote transparency on cancer reporting by hospitals in Arizona.
- The registry was invited to present on the ACR reporting objective at the American Cancer Society Cancer Action Network Summit. At the summit, the registry engaged with partners on how to tackle this issue of delinquent reporting. Summit attendees appreciated the transparency and sharing the effects of delinquent cancer reporting on cancer surveillance.

Concluding Remarks

The decision to include an objective on cancer reporting in the statewide cancer plan highlights the issue's importance. By elevating the concern, we are informing all delinquent facilities that corrective action is needed. The next steps are to continue monitoring case reporting by mandated facilities, and to work with members of the Arizona Cancer Coalition Policy work group to implement strategies proposed in the cancer plan. These strategies are designed to be realistic, attainable, and sustainable within 5 years.

The registry will create materials that share data with hospital systems on the status of their case reporting. We will reward those that meet their targets consistently with a certificate or acknowledgement on our website to show which hospital systems are following cancer reporting guidelines. Improving the timeliness of cancer reporting to the registry will improve the completeness and accuracy of cancer registry data to inform research, treatment, and early detection of cancer.

Registry contact: Arizona Central Cancer Registry (**https://www.azdhs.gov/policy-intergovernmental-affairs/cancer-registry/index.php**); Link to the cancer dashboard (**https://www.azdhs.gov/policy-intergovernmental-affairs/cancer-registry/index.php#data-dashboard)**

This content was originally presented as a poster at NCRA's 51st Annual Educational Conference, Orlando, Florida, May 2025.

Using eMaRC Lite to Streamline Pathology Report Reviews and Enhance Rapid-Case Ascertainment Studies

Lauren Maniscalco, Brent Mumphrey, Meichin Hsieh, Xiao-Cheng Wu

Summary

The Louisiana Tumor Registry (LTR)'s new Reportability application programming interface (API) increased the number of false-positive reports by 30%, resulting in a significant backlog of e-path reports that require manual review. This backlog hindered rapid case ascertainment for numerous ongoing research studies and timely identification of reportable cancer cases. To reduce false-positive e-path reports, we implemented the National Program of Cancer Registries' (NPCR's) eMaRC Lite software to supplement the Reportability API. eMaRC Lite significantly reduced the false-positive e-path reports, lowering our pathology report screeners' workload. LTR developed more eMaRC Lite models to help identify potentially eligible cases for our rapid case ascertainment studies.

Challenges

- In late 2023, the license for the pathology screening software used by the laboratories reporting to the LTR expired and was not renewed. This software was installed at the pathology laboratories and transmitted potentially reportable e-path reports to LTR. After the license expired, we asked all e-path laboratories to transmit e-path reports in HL7 format, for which we implemented a new API that uses natural language processing to identify potentially reportable e-path reports for manual review.
- Due to the lack of maturity of the new API, the number of false-positive e-path reports increased by 30%. The increased manual review workload slowed down the screening process, leading to an extensive backlog of e-path reports.
- The delay in screening e-path reports to determine reportability adversely affected LTR's rapid case ascertainment for numerous ongoing studies and timely data reporting.

Solutions

Reducing non-reportable pathology reports in the registry database:

- » Step 1. Filter pathology reports through the Reportability API; those deemed reportable become screening tasks in the registry database.
- » Step 2. Export screening tasks and run through the LTR-specific configuration of NPCR's eMaRC Lite.
- » Step 3. For reports eMaRC Lite deems nonreportable, create a mass change to code them as non-reportable in the registry database.
- » Step 4: Remaining e-path screening tasks are reviewed by pathology screeners.

Identifying cases for rapid-case ascertainment studies:

- » Step 1. Reports filtered by the Reportability API and eMaRC Lite are exported and scanned by a project-specific eMaRC Lite configuration.
- » Step 2. Reports deemed potentially eligible for a particular project are added to the "Special Study" in the registry database.
- » Step 3. The study coordinator reviews the potentially eligible reports for inclusion based on the eligibility criteria.

Results

Reducing nonreportable pathology reports in the registry database:

- » The eMaRC Lite software reduced the number of e-path screening tasks by 45%.
- » The false-negative rate is about 1%.

Identifying cases for rapid-case ascertainment studies:

- » eMaRC Lite software has been configured for two special projects: one for collecting pre-invasive cervical cancers and one for identifying colon cancer cases prior to initiation of chemotherapy.
- » About 75% of the cases identified for the pre-invasive cervical cancer project are reportable.

Concluding Remarks

- eMaRC Lite has been invaluable to our registry in terms of reducing false-positive e-path reports and saving staff time.
- The ability to tailor the eMaRC Lite for special projects enables us to continue moving forward with our commitment to cancer research.

Registry contact: Louisiana Tumor Registry (**https://sph.lsuhsc.edu/louisiana-tumor-registry/**); Link to Louisiana Interactive Data Visualizations (**https://sph.lsuhsc.edu/louisiana-tumor-registry/data-usestatistics/louisiana-data-interactive-statistics/**)

This content was originally presented as a poster at NCRA's 51st Annual Educational Conference, Orlando, Florida, May 2025.

Veterans' Data Submissions Improve High-Quality Cancer Data in Maine and Address Historic Gaps

Kathy Boris, Jackie Neas, Carolyn Bancroft, Kim Haggan

Summary

In September 2024, after 2 years of dialogue between the Maine Cancer Registry (MCR) and the Veterans Administration (VA), MCR received a data submission from the VA National Oncology Program for the first time in more than 20 years. Data submissions included diagnosis years 2007 through 2022. A total of 2,667 records were submitted to Maine. As a result of the data submissions, more than 700 new tumor records were added and 144 death certificate only (DCO) cases were updated in the MCR database.

Challenge

Prior to 2024, MCR had not received data from Togus (our local VA facility) since 2006. While MCR and the VA had established a standing order in 2022, efforts to finalize a data sharing agreement stalled between 2022 and 2023, and no data were transferred.

Based on historic data submissions from Togus, MCR estimated that 400 to 500 cases per year were missed due to lack of VA data.

Solution

After establishing a standing order in 2022, the MCR director met with members of the VA National Oncology Program at North American Association of Central Cancer Registries (NAACCR) summer forums 2023 and 2024 to discuss efforts to resume receiving VA data submissions. These efforts were supported by federal legislation passed in 2024 that required the VA to begin working with central cancer registries to report VA data.

In the fall of 2024, MCR followed up with leadership at the VA National Oncology Program. While they were unable to access the database at Togus, they were able to transfer data for Maine residents from the VA's national central cancer registry.

Results

MCR received 2,667 cases for diagnosis years 2007 to 2022 from the central VA. Although the data are likely incomplete due to lack of certified oncology data specialist (ODS-C) staff in the Togus cancer registry, this transfer is more data than we have received from the VA in nearly 20 years. As a result of the data submission, we updated 144 DCO cases and added 704 new tumor records to our database prior to our 2024 NPCR data submission (see Figure 1). More than 1,200 cases remain for consolidation that may or may not update treatment information in our database.

The VA submissions make our data more complete and higher quality than they have been in over a decade, especially for veterans.

Concluding Remarks

We will continue to partner with the VA to ensure that data transfers occur—and hopefully increase— to address any backlog and delays in cancer reporting due to staffing shortages. MCR still has not established contact with Togus and will continue working with national VA partners to ensure that cancer data for veterans in Maine are shared. We are also exploring ways to share data from MCR with the VA National Oncology Program.

Registry contact: Maine Cancer Registry (**https://www.maine.gov/dhhs/mecdc/data-reports/diseases/chronic-disease/cancer/cancer-registry**); Link to Maine Cancer Registry Available Reports (**https://www.maine.gov/dhhs/mecdc/data-reports/diseases/chronic-disease/cancer/cancer-registry/cancer-registry-reports**)

This content was originally presented as a poster at NCRA's 51st Annual Educational Conference, Orlando, Florida, May 2025.

2024 NPCR Maryland Success Story

Improving Cancer Reporting Compliance Across Maryland

Tyler Adamson, MPH, Epidemiology Team Manager

Summary

Maryland Cancer Registry (MCR) staff asked the Maryland Board of Physicians to draft a letter explaining the consequences of delinquent cancer case reporting. Since there is no formal enforcement mechanism within the statute, the MCR used a rehabilitative and collaborative approach to work with delinquent facilities to get reporting up-to-date and build facility capacity to ensure reporting compliance. The solution was time intensive but allowed for tailored support of the facilities and opportunities to use data modernization initiative (DMI) approaches to support registry infrastructure and capacity. The MCR also introduced new case completeness awards to encourage and reward compliance. These efforts laid the groundwork for working with facilities across the state to improve reporting, expand DMI, celebrate success, and enhance data collection and completeness.

Challenges

- Several facilities in Maryland were delinquent in reporting their cancer cases to the MCR.
- Two facilities with significant case counts were behind in reporting to the MCR.
- No formal enforcement mechanism exists for late or non-reporting in MCR law or regulations.
- Follow-up communication with delinquent facilities was time-consuming and unproductive.
- Compliance issues were complicated by some facilities' lack of responsiveness, buy-in, and capacity.

Solutions

- The MCR worked with the Maryland Board of Physicians and Maryland Department of Health (MDH) Deputy Secretary for Public Health Services to draft a letter outlining possible consequences of noncompliance with MCR law. The letter was sent to two delinquent facilities requesting corrective action plans.
- MCR staff reviewed the corrective action plans submitted by the facilities.
- Staff from the MCR and the MCR's quality assurance and data management contractor held biweekly check-ins with the two delinquent facilities to review case reporting and discuss ongoing capacity-related challenges and solutions.
- The MCR encouraged the use of DMI activities for delinquent facilities to enhance registry capacity. For example, one of the two facilities began piloting the use of Fast Healthcare Interoperability Resources (FHIR) to streamline and automate some aspects of the reporting process.
- To honor the MCR's late program manager, Kimberly Stern, MCR staff developed the Kimberly Stern Case Completeness Awards. Given out annually, the awards highlight facilities in the state that met three case completeness criteria: at least 90% case completeness (submitted within 9 months), at least 90% reconciliation of disease index review (reconciled and submitted within 90 days), and at least 90% reconciliation of death follow back review (reconciled and submitted within 90 days). The awardees receive copies of their certificates and are highlighted on various channels, including newsletters, email blasts, and at Tumor Registrars Association of Maryland (TRAM) meetings.

Results

- Facilities on remediation plans made progress toward compliance.
- The MCR made incremental progress year-to-year, culminating in more than 90% of case reporting for 12-month data based on a preliminary analysis of data submitted in 2024.
- The MCR used DMI activities to help delinquent facilities improve.
- The MCR awarded eight facilities with the 2023 Kimberly Stern Case Completeness Award. These certificates were given to recipients in 2024.

Sustaining Success

- The MCR will maintain regular check-ins with facilities to ensure ongoing compliance.
- The MCR will continue to monitor and track case reporting.
- The MCR will use collaborative processes and DMI approaches for other facilities that struggle with timely reporting.

Registry contact: Maryland Cancer Registry (**https://health.maryland.gov/phpa/cancer/Pages/mcr_home.aspx**); Link to Maryland Surveillance Data and Reports (**https://health.maryland.gov/phpa/cancer/Pages/surv_data-reports.aspx**)

This content was originally presented as a poster at NCRA's 51st Annual Educational Conference, Orlando, Florida, May 2025.

Incorporating 2007–2022 Veterans Administration Data into the Cancer Registry in Less than One Month

Deirdre Rogers, Ph.D., ODS, Director, Mississippi Cancer Registry

Summary

In response to language in the federal FY 2023 Appropriations Bill, the Central Veterans

Administration (VA) Cancer Registry provided the Mississippi Cancer Registry (MCR) with more than 11,000 cancer cases for Mississippi residents in the VA Cancer Registry. Unfortunately, this file was transferred on October 31, 2024, and the National Program of Cancer Registries (NPCR) annual data submission deadline was November 30, 2024.

Including these cases in the updated cancer statistics required registry staff to develop a process to incorporate them in the registry database quickly, efficiently, and accurately. The resulting plan was successful.

Challenges

- The Coastal VA Hospital had not reported cases to the MCR since 2009 since they were not required to follow the state cancer reporting law. This affects the cancer incidence rates in the Coastal area of the state.
- A process had to be developed to share cases from the Central VA System with state registries.
- State cancer registries were instructed to work with the local VA hospitals to get the cases, but the hospital on the Mississippi Coast did not have a cancer registry contact.
- On October 31, 2024, the Central VA Cancer Registry transferred more than 11,000 cases to the MCR for any Mississippi resident in the Central VA Cancer Registry System. MCR needed to submit their annual cancer data to NPCR by November 30, 2024.

Solution

The MCR used linkages to eliminate cases that had already been submitted by the Coastal VA. This reduced the volume of cases. Modified steps were developed to process the cases by new patients, new primary cancers, or new reports on an existing primary cancer using CDC's Registry Plus software.

Results

The MCR staff were able to complete modified data processing, and the data were incorporated in the November 2024 submission file. The following steps were developed to process the cases:

1. The MCR director linked the Coastal VA hospital cases using probabilistic matching to eliminate cases the MCR had received from that facility in past submissions. This reduced the number to about 4,000 cases.
2. Edits were corrected in each file, and common quality issues were reviewed in the cases and corrected.
3. The files of cases were loaded in CDC's Registry Plus database system, which conducted some automated linkage of patients and tumors. More than 2,400 cases still needed to be processed to determine new patients and primaries and to consolidate this new report with prior reports from other facilities.
4. The MCR director and manager decided to complete all cases for diagnosis year 2022. For other years, new patients and primary tumors would be processed completely. If a cancer case could replace an existing death certificate only case, the case would be completed. Lastly, for cases already in CRS Plus software, staff would only process cases that would cause a change in birthdate, race, county of residence, diagnosis date, primary site, histology, or behavior.
5. All certified staff participated in processing these cases. We learned this process would work when we have a large data submission that must be completed in a short time.

Concluding Remarks

Next steps include completing the record processing prior to November 2025. Incorporating these cases was successful in reducing the death certificate only cases for all years involved and will provide more accurate incidence rates, especially for the Coastal region of Mississippi.

Registry contact: Mississippi Cancer Registry (**https://www.umc.edu/cancerinstitute/Cancer-Research/Cancer-Registries/Mississippi%20Cancer%20Registry/HOME-PAGE-CONTENT.html**); Link to Mississippi Cancer Facts and Stats (**https://www.cancer-rates.com/ms/**)

This content was originally presented as a poster at NCRA's 51st Annual Educational Conference, Orlando, Florida, May 2025.

Monitoring and Evaluating Hospital Reporting Timeliness to Improve Registry 12-Month Completeness

Kaitlin R. Kruger, MS (Data Administration Manager) and Emily C. Bunt, MA (Registry Manager)

Summary

During a townhall meeting last year, CDC's National Program of Cancer Registries (NPCR) shared that they were considering reinstating the Advanced National Data Quality standard that would require 12-month data to be 90% complete. To meet this requirement, Ohio Cancer Incidence Surveillance System (OCISS) developed a new data quality report to monitor and evaluate hospital reporting completeness and timeliness.

Challenge

- In accordance with Ohio laws and rules (Ohio Administrative Code, https://codes.ohio.gov/ohio-administrative-code/rule-3701-4-02) cancer must be reported to OCISS within 6 months of date of diagnosis or first contact with the facility.
- Most Ohio reporting facilities do not meet the 6-month reporting timeline for many reasons, including staffing, funding, and treatment data requirements.
- Per NPCR program standards (https://www.cdc.gov/national-program-cancer-registries/about/npcr-standards.html) 24-month data submitted to NPCR must be at least 95% complete and 12-month data are encouraged to be 90% complete.
- Since 2001, Ohio has met the 90% threshold for 12-month data only once, in 2018.

Solution

- The vast majority of Ohio's overall volume is reported by hospitals.
- OCISS developed a method to monitor timeliness and completeness of hospital submissions.
- This report displays the completeness and timeliness of the 12-month data that will be submitted to NPCR in November.
- The timeliness report helps identify hospitals that are not reporting on schedule for follow-up.

Results

OCISS created the timeliness report by:

1. Determining report structure and content.
2. Developing queries and a spreadsheet template for generating the report.
3. Executing queries, calculating percentages, and formatting the report each quarter.
4. Sharing reports with hospitals.
5. Comparing reports for each quarter and reviewing results.

Ohio's advisory committee—comprised of hospital registries, local health departments, and researchers—provided feedback on the development of the report.

Figure 1. Ohio's 2023 Cancer Data Reporting Records Submitted within 6 Months

RSID	Total DxYear 2023 Received	Total Expected	% expected records received	% of reporting year elapsed	On track?	% of A records submitted w/in 6 months of date 1st contact
58	23	46	50%	75%	-25%	0%
59	90	129	70%	75%	-5%	11%
60	0	4	0%	75%	-75%	0%
61	250	283	88%	75%	13%	77%
62	320	476	67%	75%	-8%	15%
65	0	655	0%	75%	-75%	0%
66	10	0	0%	75%	-75%	0%
67	121	216	56%	75%	-19%	4%
70	339	391	87%	75%	12%	71%
71	1309	1258	104%	75%	29%	73%
72	6198	7595	82%	75%	7%	59%

Registry contact: Ohio Cancer Incidence Surveillance System (OCCISS) (**https://odh.ohio.gov/know-our-programs/ohio-cancer-incidence-surveillance-system/welcome-to/**); Link to Ohio Cancer Data and Statistics (**https://odh.ohio.gov/know-our-programs/ohio-cancer-incidence-surveillance-system/data-statistics/data-statistics**)

This content was originally presented as a poster at NCRA's 51st Annual Educational Conference, Orlando, Florida, May 2025

Ohio's data submission results:

- » For data submission year 2021, Ohio's 12-month data were 74.02% complete.
- » For data submission year 2022, Ohio's 12-month data were 81.37% complete.
- » For data submission year 2023, Ohio's 12-month data were 80.32% complete.
- » Data submission results for 2024 are not yet available. However, Ohio submitted an additional 1,200 cases for 12-month data this year, which is about a 2% increase for the NPCR incidence count.

Lessons learned:

- » Use a larger year range for calculating the average expected number of cases.
- » Provide additional definitions for each column in the report using comment functionality.

Concluding Remarks

- The timeliness report provides consistency, efficiency and transparency in monitoring, evaluating, and improving hospital reporting timeliness.
- By creating a template and queries, this is sustainable and is now incorporated into our standard operating procedures.
- The new report encourages accountability for the central registry and hospitals. It holds ourselves accountable to follow up with facilities that do not meet reporting requirements. It holds our hospitals accountable to improve their reporting timeliness.
- To take this a step further, we plan to create a certificate or award for hospitals that are consistently performing well. We have also considered expanding the report to include non-hospital facilities which make up the next highest reporting volume. Additionally, our goal is to complete the v25 upgrade timelier in 2025.

2024 NPCR Oklahoma Success Story

Improving Reporting Timeliness Compliance in Oklahoma

Meagan Carter, MS and Christy Dabbs, ODS-C

Summary

The Oklahoma Central Cancer Registry (OCCR) has struggled with Oklahoma facilities reporting timely data. In addition to generating quarterly compliance reports and adjusting registry communication strategies, OCCR implemented remediation plans for facilities that were chronically delinquent in reporting cases.

Challenges

- Many Oklahoma reporting facilities have been behind on reporting cancer cases to OCCR due to high staff turnover rates, insufficient resources, competing priorities, and lack of understanding of cancer case reporting mandates.
- Facility administrators may not understand the steps required to comply with reporting requirements.

Solutions

- OCCR emailed quarterly compliance reports to facility supervisors and mailed letters to facility administrators informing them of their compliance status for the current and previous reporting years.
- OCCR implemented a remediation plan for chronically delinquent facilities to catch them up on reporting and maintain compliance. OCCR staff met with facility administrators and supervisors to discuss expectations for the remediation plan and set final reporting deadlines.
- Remediation plans required facilities to report cases to OCCR twice a month instead of monthly. All cases completed in that period must be reported on the 15th and 30th of each month.
- The OCCR program director and data manager reviewed file submissions the day after the deadlines. They contacted facilities that did not submit a file.
- When facilities maintained compliance for 3 consecutive months, they were released from the bimonthly reporting requirement.

Results

- OCCR's compliance efforts resulted in 20,997 cases being reported for preliminary 2023 data at annual data submission, compared to 19,662 cases reported to NPCR for 2022 preliminary data, demonstrating a 7% increase.
- In July 2024, 46% of facilities were fully compliant for 2022 and 2023 reporting, with 36% facilities compliant for only 2022 and 5% compliant for only 2023. For first quarter (Q1) 2024, 39% were fully compliant for 2023 and Q1 2024, with 33% compliant for only 2023 and 12% compliant for only Q1 2024.
- During the remediation plan piloting phase in summer 2023, one hospital and one ambulatory surgery center were placed on remediation plans.
- The hospital immediately acquired contract services to start reporting within 3 months of the start of the plan and were caught up on reporting within 6 months.
- The ambulatory surgery center trained a part-time employee to report their cases. As of December 2024, the center was still working to catch up on the backlog that was left from previous staff vacancies.
- As of December 2024, 12 facilities are on remediation plans.

Concluding Remarks

- Sending quarterly compliance communication to facility supervisors and administrators increased the overall number of cases that OCCR reported for preliminary 2023.
- Remediation plans improve communication with facilities and provide an opportunity for facility leaders and OCCR to monitor reporting more closely.

Registry contact: Oklahoma Central Cancer Registry (**https://oklahoma.gov/health/health-education/data-and-statistics/center-for-health-statistics/oklahoma-central-cancer-registry-occr.html**); Link to Oklahoma OK2Share data (**https://www.health.state.ok.us/ok2share/index.shtml**)

This content was originally presented as a poster at NCRA's 51st Annual Educational Conference, Orlando, Florida, May 2025.

Journal of Registry Management Continuing Education Quiz—SUMMER 2025

INFERRING UNKNOWN RACE IN CENTRAL CANCER REGISTRIES

Cari Vida, RHIA, ODS-C, Contributing Editor

This quiz is derived from the article, "Inferring Unknown Race in Central Cancer Registries" by Francis P. Boscoe, PhD.

After reading the article and completing the quiz, participants will be able to:

- Understand the completeness of race information for data certification among US Central Cancer Registries.
- Describe the three methods assessed for imputing missing race information for cancer patients.
- Understand the widely used Bayesian Improved Surname Geocoding (BISG) method.

1. What state central registry had the highest percentage of race missing for cancers diagnosed from 2017–2021?
   a) Arizona
   b) Oregon
   c) Washington
   d) California
2. Patient race is considered missing (code 99 unknown) in what percentage of cancer patients in the US?
   a) 3.4%
   b) 1.8%
   c) 1.0%
   d) 3.0%
3. What is the Central Cancer Registry National "gold" and "silver" standard for missing patient race?
   a) 3% for gold and 5% for silver
   b) 1% for gold and 3% for silver
   c) 5% for gold and 7% for silver
   d) 2% for gold and 4% for silver
4. Supplementing names lists with what factor has proved to yield complete and accurate inferences about race and ethnicity?
   a) Patient place of birth
   b) Patient address information
   c) Patient socioeconomic status
   d) Patient political party
5. Which of the following is an available R software package that uses available name lists and census data on the racial and ethnic composition of small geographic areas to make individual predictions about race and ethnicity?
   a) BISG
   b) API
   c) NHIA
   d) WRU
6. Bayesian, in the acronym BISG, refers to what rule to find the conditional probability that an individual belongs to a given race/ethnicity given their name, residential location, and other demographic variables?
   a) Bayes
   b) Positive Predictive Value
   c) Cohen's Kappa
   d) Proportion
7. Per Table 2, what is the overall accuracy for the method of census surnames + voter roll names with an 85% threshold?
   a) 92.1%
   b) 96.0%
   c) 95.0%
   d) 97.4%
8. Per the paper, which of the following is NOT a suggested way that the BISG method could be improved?
   a) Incorporating subcounty-level geocoding
   b) Including age and sex
   c) Expanding to identify American Indians and Alaska Natives
   d) Including marital status
9. What is the reason that name-only methods are good at identifying Hispanic and API individuals but not Black individuals, per the paper?
   a) Hispanic and API names are more common.
   b) Hispanic and API names have more overlap with non-Hispanic White names than Black names.
   c) Hispanic and API names have less overlap with non-Hispanic White names than Black names.
   d) Black names have less overlap with non-Hispanic White names than Hispanic and API names.
10. Per the paper's example, the name Martin has what distribution rate?
    a) 6% Hispanic
    b) 85% White
    c) 25% Black
    d) 5% AIAN

Purchase Quiz to Earn CE:

1. Go to **http://www.cancerregistryeducation.org/jrm-quizzes**
2. Select quiz and "Add to Cart" (You may be prompted to login using your NCRA login).
3. Continue through the checkout process.
4. Once purchase is complete, the quiz will load automatically into "My Learning Activities" page.

National Cancer Registrars Association
CALL FOR PAPERS

The *Journal of Registry Management*, official journal of the National Cancer Registrars Association (NCRA), announces a call for original manuscripts on registry methodology or research findings related to the 7 subjects listed below and related topics.

Topics:

1. **Birth Defects Registries**
2. **Cancer Registries**
   a. **AJCC TNM Stage**
   b. **Cancer and Socioeconomic Status**
   c. **Cancer and Health Disparities**
3. **Trauma Registries**
4. **Recruitment, Training, and Retention**
5. **Public Relations**
6. **Quality Review**
7. **Registry Management**

Contributed manuscripts are peer-reviewed prior to publication. Manuscripts of the following types may be submitted for publication:

1. **Methodology Articles** addressing topics of broad interest and appeal to the readership, including methodological aspects of registry organization and operation.
2. **Research articles** reporting findings of original, reviewed, data-based research.
3. **Primers** providing basic and comprehensive tutorials on relevant subjects.
4. **"How I Do It" Articles** describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. The "*How I Do It*" feature in the *Journal* provides registrars with an informal forum for sharing strategies with colleagues in all types of registries.
5. **Opinion papers/editorials** including position papers, commentaries, essays, and interviews that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management.
6. **Bibliographies** which are specifically targeted and of significant interest will be considered.
7. **Letters to the Editor** are also invited.

Address all manuscripts to: Nadine Walker, MS, ODS-C, Editor-in-Chief, *Journal of Registry Management,*
(703) 299-6640 ext. 327, JRMEditor@ncra-usa.org.

Manuscript submission requirements are given in "Information for Authors" found near the back of each *Journal* and on the NCRA website at **https://www.ncra-usa.org/About/Publications/Journal-of-Registry-Management**.

INFORMATION FOR AUTHORS

The Journal of Registry Management, the official journal of the National Cancer Registrars Association (NCRA), invites submission of original manuscripts on topics related to management of disease registries and the collection, management, and use of cancer, trauma, birth defects, HIV/AIDS, and other disease registry data. JRM is a peer-reviewed, open-access, online-only journal and is published quarterly.

JRM encourages authorship by registrars who are ODS-certified (Oncology Data Specialist); special value is placed on manuscripts with ODS-certified professionals' collaboration and publication of articles or documents related to the registry profession. Three NCRA continuing education (CE) credits are awarded for published articles or documents, and additional information can be found at the following URL: **https://www.ncra-usa.org/ODS-Credential/Current-ODS/Submit-CEs/CE-Eligible-Activities**

Manuscripts may be submitted for publication in the following categories:

- Articles addressing topics of broad interest and appeal to the readership, including methodology papers about registry organization and operation
- Research papers reporting findings of original research, literature reviews, data-based research
- Primers providing tutorials on relevant subjects
- "How I Do It" papers

JRM invites submission of:

- Opinion papers/editorials including position papers, commentaries, and essays that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management
- Letters to the Editor
- Specifically targeted bibliographies of significant interest

Previously published material will be considered for publication only when it is of special and immediate interest to the readership.

Please submit manuscripts and articles here: https://srvy.pro/2CXB3FV.

Manuscript Preparation Guidelines

The following guidelines are provided to help prospective authors prepare manuscripts for the JRM and facilitate technical processing of submissions. Failure to follow the guidelines may delay consideration of your manuscript.

Authors who are unfamiliar with preparation and submission of manuscripts for publication are encouraged to contact the Editor for clarification or additional assistance. All correspondence and questions about manuscripts should be sent to **JRMEditor@ncra-usa.org**. Telephone inquiries may be directed to (703) 299-6640 ext. 327

Cover Letter and Signature Page

An accompanying cover letter should include the name, mailing address, email address, and telephone number of the corresponding author(s).

An authors' signature page can be a scanned copy containing all the author's signatures, or an email acknowledgement from each author can be sent to **JRMEditor@ncra-usa.org**. See Copyright section below for instructions on authors' permissions.

Manuscript Types

The terms *manuscripts, articles,* and *papers* are used synonymously herein. Number the manuscript pages consecutively with the title page as page 1, followed by the abstract, text, references, and visuals.

Articles

Articles should follow the standard format for research reporting (Introduction, Methods, Results, Discussion, References). The introduction will normally include background information, and a rationale/justification as to why the subject matter is of interest. The discussion often includes a conclusion subsection. Comprehensive references are encouraged, as are an appropriate combination of tables and figures (graphs).

Methodology/Process Papers

Methodology papers should follow the standard format for research reporting (Introduction, Methods, Results, Discussion), or for explanatory papers not reporting results (Introduction, Methods, Discussion).

"How I Do It" Articles

The "How I Do It" feature in the *JRM* provides registrars with a forum for sharing strategies with colleagues in all types of registries. These articles describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. When shared, these innovations can help registry professionals improve their skills, enhance registry operations, or increase efficiency.

"How I Do It" articles should be 1,500 words or less (excluding references) and can contain up to 2 tables or figures. To the extent possible, the standard headings (Introduction, Methods, Results, Discussion) should be used. If results are not presented, that section may be omitted. Authors should describe the problem or issue, their solution, advantages (and disadvantages) to the suggested approach, and their conclusion. All submitted "How I Do It" articles will have the benefit of peer/editorial review.

Manuscript Format and Structure

Authors

Each author 's name, degrees, certifications, title, professional affiliation, and email address must be noted on the title page exactly as it is to appear in publication. The corresponding author should be noted, with mailing address included. Joint authors should be listed in order of their contribution to the work. A maximum of 6 authors for each article will be listed. Consider using a "working group" or committee name if there are more than 6 authors and list the lead author(s).

Title

Authors are urged to choose a title that accurately describes the manuscript's content. Every effort will be made to use the title as submitted; however, *JRM* reserves the right to select a title that is consistent with editorial and production requirements.

Abstract

A brief abstract must accompany each article or manuscript (word limit, 350). The abstract should summarize the main point(s) and quickly give the reader an understanding of the manuscript's content. It should be placed on a page by itself, immediately following the title page.

Key words

Authors are asked to provide up to 5 alphabetized key words or phrases which will be used in compiling the Annual Subject Index. Key words should be included directly under the abstract on the abstract stand-alone page.

Length

The word count for manuscripts should not exceed 3,000 words. Word count guidelines exclude abstracts, figure legends, and table notes. Authors are invited to contact the Editor regarding submissions of markedly longer manuscripts.

Style

Prepare manuscripts using the *American Medical Association Manual of Style,* 11th edition (2020). All sections of the paper should be single-spaced. Double-space between paragraphs and sections.

Visuals

Use visuals selectively to supplement the text. Visual elements—charts, graphs, tables, diagrams, and figures—will be reproduced exactly as received. Copies must be clear and properly identified. Each visual must have a brief, self-explanatory title. Submit each visual on a separately numbered page at the end of the manuscript, following the references. Visuals for research papers, articles, and methodology/process papers are limited to 5; Visuals for "How I Do It" papers and all other submissions are limited to 2.

Attribution

Authors are to provide appropriate acknowledgment of products, activities, and support, especially for those articles based on or utilizing registry data (including acknowledgment of hospital and central registrars). Appropriate attribution is also to be provided to acknowledge federal funding sources of registries from which the data are obtained.

References

References should be carefully selected, and relevant. References must be numbered in order of their appearance in the text. At the end of the manuscript, list the references in the order they are cited; do not list references alphabetically. Journal citations should include author, title, journal, year, volume, issue, and pages. Book citations should include author, title, city, publisher, year, and pages. Authors are responsible for the accuracy of all references.

Examples:

1. LeMaster PL, Connell CM. Health education interventions among Native Americans: a review and analysis. *Health Educ.* 1995;21(4):521-538. doi:10.1177/109019819402100413
2. Hanks GE, Myers CE, Scardino PT. Cancer of the prostate. In: DeVita VT, Hellman S, Rosenberg SA. *Cancer: Principles and Practice of Oncology.* 4th ed. J.B. Lippincott Co.; 1993:1073-1113.

Affirmations

Copyright

Authors submitting a manuscript do so on the understanding that if it is accepted for publication, copyright of the article, including the right to reproduce the article in all forms and media, shall be assigned exclusively to NCRA. NCRA will not refuse any reasonable requests by the author(s) for permission to reproduce any of his or her contributions to the *JRM.*

Authors Permissions

Further, the manuscript's accompanying cover letter, signed by all authors, must include the following statement: "We, the undersigned, transfer to the National Cancer Registrars Association, the copyright for this manuscript in the event that it is published in *Journal of Registry Management*." Failure to provide the statement will delay consideration of the manuscript. The authors' signature page can be a scanned copy containing all corresponding authors signature, or an email acknowledgement from each author can be sent to JRMEditor@ncra-usa.org. It is the author's responsibility to obtain necessary permission when using material (including graphs, charts, pictures, etc) that has appeared in other published works.

Originality

Articles are reviewed for publication assuming they have not been accepted or published previously and are not under simultaneous consideration for publication elsewhere. If the article has been previously published or significantly distributed, this is to be noted in the cover letter for consideration.

Editing

Journal of Registry Management reserves the right to edit all contributions for clarity and length. Minor changes (punctuation, spelling, grammar, syntax) will be made at the editorial staff's discretion. Substantive changes will be verified with the author(s) prior to publication.

Peer Review

JRM follows a double-blind peer review process. Contributed manuscripts are reviewed prior to publication, generally by 2 reviewers. The Journal Editor makes the final decision regarding acceptance of manuscripts. Receipt of manuscripts will be acknowledged promptly, and corresponding authors will be advised of the status of their submission as soon as possible.

Ethics

Conflict of Interest

As part of the online submission process, corresponding authors must confirm if they or their coauthors have conflicts of interest to declare, and to provide details of these. These include all financial and non-financial interests and relationships, direct or indirect, or other situations that might raise questions of bias in the work reported or the conclusions, implications, or opinions stated. Authors should also disclose any conflict of interest that may have influenced either the conduct or the presentation of the research to the editors, including but not limited to, close relationships with those who might be helped or hurt by the publication, academic interests, and rivalries, and any personal, religious, or political convictions relevant to the topic at hand. If the manuscript is published, conflict of interest information will be communicated in a statement within the published paper. If any reviewer believes there is likely to be a perception of a conflict of interest in relation to their review of a submitted manuscript, they will notify the Editor-in-Chief. The review will then be assigned to other Editorial Advisory Board (EAB) members. Authors should identify individuals who provide writing assistance and disclose the funding source for this assistance.

Informed Consent

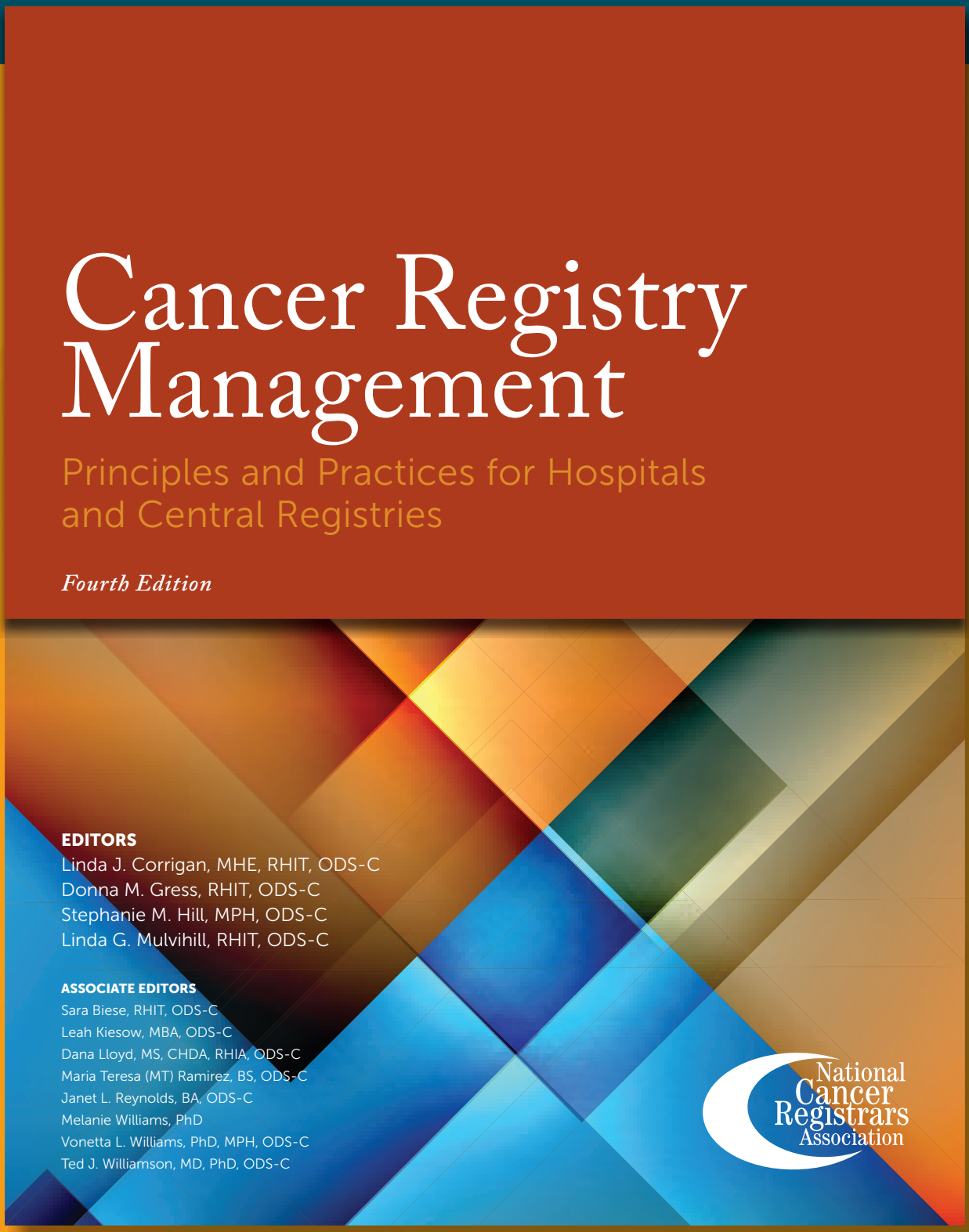
Individuals have a right to privacy that should not be infringed without informed consent. Identifying information, including patients' names, initials, or hospital numbers, should not be published in written descriptions, photographs, and pedigrees, unless the information is essential for scientific purposes and the patient (or parent or guardian) gives written informed consent for publication. Informed consent for this purpose requires that a patient who is identifiable be shown the manuscript to be published. Identifying details should be omitted if they are not essential. Complete anonymity is difficult to achieve; however, an informed consent should be obtained if there is any doubt. For example, masking the eye region in photographs of patients is inadequate protection of anonymity. If identifying characteristics are altered to protect anonymity, such as in genetic pedigrees, authors should provide assurance that alterations do not distort scientific meaning and editors should so note.

Human and Animal Rights

Research involving human subjects, including identifiable human material or data, must have been performed in accordance with the Declaration of Helsinki and must have been approved by an appropriate ethics committee. A statement detailing this, including the name of the ethics committee and the reference number where appropriate, must appear in all manuscripts reporting such research. Although *JRM* does not publish animal research, it may accept it in specific situations (eg, when an animal experiment is also part of a human trial). Authors interested in submitting animal research should contact the Editor-in-Chief. Any study using animals needs to state the Institutional Animal Care approval and number. Any other ethics approvals should also be listed. If no ethical approvals were required, please state this.

A definitive update of the premier textbook and desk reference for the cancer registry profession.

Order print or e-book at
www.ncra-usa.org/4thEdition *today!*



Journal of Registry Management

NCRA Executive Office
1330 Braddock Place
#520
Alexandria, VA 22314

PRESORTED
STANDARD
U.S. POSTAGE
PAID
DULLES, VA
PERMIT #6418

ADDRESS SERVICE REQUESTED

SUSTAINABLE FORESTRY INITIATIVE
Certified Sourcing
www.sfiprogram.org
SFI-00359

100% green power wind & solar

Printed on SFI fiber sourcing paper with non-petroleum, vegetable based inks and manufactured with renewable electricity