Cancer Care Delivery Research and the National Cancer Institute SEER Program
Challenges and Opportunities

Steven J. Katz, MD, MPH

OBJECTIVE To determine the proportion of breast cancers that were identified at an early stage (stage I) in different racial/ethnic groups and whether ethnic differences may be better explained by early detection or by intrinsic biological differences in tumor aggressiveness.

RESULTS Of 373 563 women with invasive breast cancer, 268 675 (71.9%) were non-Hispanic white; 34 928 (9.4%), Hispanic white; 38 751 (10.4%), black; 25 211 (6.7%), Asian; and 5998 (1.6%), other ethnicities. Mean follow-up time was 40.6 months (median, 38 months). Compared with non-Hispanic white women diagnosed with stage I breast cancer (50.8%), Japanese women (56.1%) were more likely to be diagnosed (OR, 1.23 [95% CI, 1.15-1.31], P < .001) and black women (37.0%) were less likely to be diagnosed (OR, 0.65 [95% CI, 0.64-0.67], P < .001). Actuarial risk of death from stage I breast cancer at 7 years was higher among black women (6.2%) than non-Hispanic white women (3.0%) (HR, 1.57 [95% CI, 1.40-1.75]; P < .001), and lower among South Asian women (1.7%) (HR, 0.48 [95% CI, 0.20-1.15]; P = .10). Black women were more likely to die of breast cancer with small-sized tumors (9.0%) than non-Hispanic white women (4.6%) (HR, 1.96 [95% CI, 1.82-2.12]; P < .001); the difference remained after adjustment for income and estrogen receptor status (HR, 1.56 [95% CI, 1.45-1.69]; P < .001).

CONCLUSIONS AND RELEVANCE Among US women diagnosed with invasive breast cancer, the likelihood of diagnosis at an early stage, and survival after stage I diagnosis, varied by race and ethnicity. Much of the difference could be statistically accounted for by intrinsic biological differences such as lymph node metastasis, distant metastasis, and triple-negative behavior of tumors.

JAMA. 2015;313(2):165-173

Initiatives to build big data to inform and improve the quality of cancer care delivery in the United States are evolving rapidly. Advances in information technology have made it increasingly more feasible to efficiently collect clinical data across facilities and institutions to address challenges related to access, delivery, and disparities in care. The National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program is uniquely positioned in the United States to leverage big data to inform oncology practice in the community. The participating regional registries comprise a population that is diverse and representative of the United States. The registries have unique state data agreements and health care facility partnerships (hospitals and pathology laboratories) that enable them to collect uniform valid clinical information about the first course of cancer treatment for virtually all patients who receive diagnoses in their regional catchment areas. These data have been used by many hundreds of investigators to generate more than 8000 scientific articles that address burden of cancer, disease etiology and behavior, disparities in cancer health outcomes, and quality of oncology care.

But there are some important limitations to SEER data. First, the level of detail about tests and treatments is limited. For example, information about adjuvant chemotherapy regimens is not routinely collected. Second, the quality and completeness of the clinical data that are collected may vary by cancer condition and across regions. Finally, information related to patient socioeconomic status is limited. For example, SEER data currently include patient-level information for race and marital status but not for income, education, or insurance status.
The SEER-Medicare data program has been a major advance because linking medical claims to SEER augmented and improved information about tests and treatments. But SEER-Medicare data lack some generalizability because they are restricted to Medicare beneficiaries and thus, for example, include less than half of patients with diagnoses of breast or colorectal cancer.

The article by Iqbal et al published in JAMA in January highlights the opportunities and challenges of using SEER data alone to address the etiology of racial and ethnic disparities in cancer presentation and survival. The authors explored the etiology of observed racial and ethnic differences in stage of diagnosis and survival. Using an appropriate design and rigorous methodology, they concluded that black-white differences in mortality are largely explained by differences in intrinsic biology of the breast cancer and health status at time of diagnosis. Their findings are consistent with several other studies that have evaluated the etiology of racial and ethnic differences in health outcomes. In addition, they considered the possible role played by cancer control programs in explaining racial and ethnic differences in disease presentation. In particular, they speculate about nonbiological factors that might explain racial and ethnic differences in the proportion of patients who receive diagnoses of early-stage disease and survival. However, considering stage or survival as an outcome of cancer control programs is problematic because of lead-time bias. The somewhat indirect approach to this question reinforces the limitations of using a disease epidemiologic framework, design, and measures to inform the effectiveness of clinical delivery programs.

A recent article by Silber et al underscores the advantage of using SEER-Medicare rather than SEER data alone to more directly address questions about the impact of cancer control strategies on outcomes in populations of patients with cancer. This innovative study examined to what extent black-white differences in survival after diagnosis of breast cancer were due to differences in treatment. The authors used a comprehensive set of valid treatment information available in the unique SEER-Medicare claims data set. They observed that differences in breast cancer treatment accounted for a very small amount of the variation in survival between blacks and whites, after elegantly controlling for black-white differences in disease presentation.

Studies using SEER program data augmented with more granular patient-reported measures directly evaluated racial and ethnic differences in the delivery of locoregional and systemic treatment for breast cancer. Innovative aspects of these studies included oversampling racial and ethnic groups, accruing patients into the studies shortly after diagnosis, and obtaining detailed information on tests and treatments by linking patient survey reports to SEER data. These studies observed no substantial black-white disparities in the initiation of clinically indicated locoregional and systemic therapies. These findings, in conjunction with the article by Silber et al, provide some reassurance regarding the level of black-white disparities in treatment and treatment-related outcomes in patients with breast cancer.

The rich and growing portfolio using SEER data reinforces the unique status of the program in population-based cancer care delivery research. The SEER program is built on a long history of partnerships between state health departments, regional cancer registries, health care facilities, and the National Cancer Institute. Indeed, no other big data initiative in oncology approaches the comprehensiveness and quality of data collection and generalizability of the results. A major challenge for SEER is to modernize the content of the data in a rapidly evolving landscape of cancer management. The first task is to leverage opportunities to obtain more granular information about rapidly emerging evaluative tests and treatments for cancer. Current data collection efforts largely depend on both passive and active reporting from pathology laboratories and hospitals. But more efficient data collection is on the horizon with advances in automated clinical data repositories and electronic medical records. For example, the SEER registries have led the way in automating the transfer and collection of pathology data across hospitals. Another emerging opportunity is SEER partnership with industry. In particular, tumor genomic and genetic testing companies may be interested in partnering with federal and state public health entities, such as SEER, to perform research that informs quality of care and the patient experience. Finally, there is a need to augment clinical and treatment information in SEER with patient-reported measures of communication, decision-making, and health outcomes. Currently, there are a number of demonstration projects that will serve as useful models for the way forward.

The SEER program has been a critically important source for information and research about cancer etiology and outcomes, patterns of treatment, and disparities in care in the community. However, other big data initiatives in oncology are growing, led by clinician groups such as the American College of Surgeons and the American Society of Clinical Oncology, as well as private industry. The SEER program will need to evolve in this rapidly changing landscape of cancer management, information technology, and partnerships to continue to thrive as a vital resource in population studies in oncology.