The Surveillance, Epidemiology, and End Results Program: A National Resource

Benjamin F. Hankey, Lynn A. Ries, and Brenda K. Edwards

Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland 20892

The fight against cancer escalated in 1971 with the introduction of the National Cancer Act during the presidency of Richard M. Nixon. Since that time, there have been dramatic developments in a number of research areas including prevention, screening, and treatment that have had a major impact on our ability to control cancer. An integral part of the fight against cancer has been the ability to monitor the occurrence of cancer in the population and assess the impact of the introduction of advances in prevention, screening, and treatment. For more than 25 years, these functions have been performed by the SEER2 Program, which is based at the NCI in the Division of Cancer Control and Population Sciences. The purpose of this review is to briefly describe cancer surveillance at the NCI. The SEER Program is the centerpiece of these activities, with various surveillance and research functions arrayed around it to take full advantage of its potential.

The SEER Program is a sequel to two earlier NCI programs, the End Results Program and the Third National Cancer Survey (1). Case ascertainment and data collection for the SEER Program began with January 1, 1973 diagnoses in several geographic areas of the United States and its territories. Those areas that have participated in the program since 1975 include the states of Connecticut, Iowa, New Mexico, Utah, and Hawaii and the metropolitan areas of Detroit, San Francisco/Oakland, Seattle, and Atlanta. Subsequent additions to the program included 10 predominantly black rural counties in Georgia in 1978 and American Indians residing in Arizona in 1980. In 1992, the program was further expanded to increase coverage of minority populations, especially Hispanics. The two new areas added were Los Angeles County and four counties in the San Jose/Monterey area south of San Francisco. Alaska natives in Alaska are currently being added to those populations covered by SEER. The SEER Program currently includes population-based data from about 14% of the United States population and is reasonably representative of subsets of the different racial/ethnic groups residing in the United States. Fig. 1 provides a map of the SEER areas, and Fig. 2 gives the percentages and sizes of various populations included in the SEER Program.

The issue of the representativeness of SEER areas with regard to cancer incidence is frequently raised. Fig. 3 presents some comparisons of interest for SEER areas versus the total United States based on ecological data from the 1990 census. SEER areas tend to be more urban and have a larger percentage of foreign-born individuals. This at least partially explains why SEER cancer mortality rates deviate somewhat from those for the total United States, even within racial/ethnic groups. However, with few exceptions, the trends in mortality rates for individual cancers observed for SEER areas are generally consistent with those for the total United States (2). These observations indicate that it is reasonable to assume that cancer incidence trends for SEER are generally representative of those for the total United States.

The objectives of the SEER Program include: (a) developing and reporting estimates of cancer incidence and mortality on a periodic basis for the total United States; (b) monitoring annual cancer incidence trends to identify unusual changes in specific forms of cancer occurring in population subgroups defined by geographic, demographic, and social characteristics and providing insight into their etiology; (c) providing continuing information on changes over time in the extent of disease at diagnosis, trends in therapy, and associated changes in patient survival; and (d) promoting studies designed to identify factors pertaining to the environment, occupation, socioeconomic status, tobacco, diet, screening practices, patterns of care, and determinants of the length and quality of patient survival that are amenable to cancer control interventions.

The primary cancer surveillance publication is the Cancer Statistics Review, which is produced annually (3). It includes data for the time period 1973 to the most recent year for which data are available. Cancer incidence, mortality, and patient survival rates and other statistics of interest are presented by anatomical site. For major cancers, the rates are also broken down by race/ethnicity and geographic area.

Monographs are also produced periodically. A monograph on the histology of cancer was published in 1995 as a supplement to the journal Cancer (4). More recently, a monograph entitled Racial/Ethnic Patterns of Cancer in the United States 1988–1992 (5) provides a description of the occurrence of the major cancers among most racial/ethnic groups in the United States. A monograph on prostate cancer entitled SEER Prostate Cancer Trends, 1973–1995 (6) that provides detailed data on incidence, mortality, patient survival, and treatment is currently being published. Other monographs are in preparation, including one on childhood cancer and one on patient survival. A detailed list of SEER publications including monographs, recent reports, data collection manuals, manuals on registry operations, and training manuals pertaining to data collection are available from the SEER Web site. It is also possible to view/print many of these publications directly from the Web.

3 The SEER Web site is located at http://www-seer.ims.nci.nih.gov/.
Data sources used by SEER include the SEER database, which contains records on more than two million cancers and is growing at the rate of more than 160,000 records per year; cancer mortality data for the total United States as obtained from the National Center for Health Statistics; and population estimates as obtained through an interagency agreement with the Census Bureau.

It is possible to obtain a SEER public use file on a CD-ROM free of charge. The file includes information on individual cancers diagnosed in SEER areas between 1973 and the most recent year for which data are available. Personal identifiers are not included. Instructions for obtaining a public use file are available at the SEER Web site. Currently, SEER sends out public use files at the rate of nearly 1500/year, up from less than 50/year 10 years ago.

Since its inception in 1973, the individual SEER cancer registries have taken advantage of the potential for research offered by such a mechanism, particularly the opportunity to conduct population-based epidemiological studies. Hundreds of studies have been conducted and reported, with support generally provided by grants from the federal government. A search of the National Library of Medicine’s database PUBMED using only the term SEER identified in excess of 570 publications for the time period 1978-1999, many of which either include an analysis of SEER data or refer to cancer statistics based on SEER data.

The SEER program contractors include a number of universities involved in cancer research. Thus, there is an infrastructure for conducting analytic epidemiological studies on a variety of emerging issues in cancer prevention and control that can be used by the NCI. The ability to do special studies was established in the early 1990s. The workscopes of SEER contracts were modified to include the capabilities of interviewing patients, conducting surveys of the covered populations, obtaining biological materials from patients and survey respondents, conducting methodological research that uses cancer registry data, and establishing tissue banks. Standard competitive procurement procedures within the SEER framework have been used to plan and fund studies on a wide range of topics that have included identification of risk factors, quality of life, statistical modeling, etiology of trends in cancer rates, and operational issues pertaining to data collection and reporting.

A wide variety of research activities use the SEER database and/or the SEER special study mechanism. An example of a study done through the SEER Special Study mechanism is a study on quality of life after treatment for localized prostate cancer that has been carried out in six SEER areas. Approximately 3500 men were interviewed shortly after diagnosis and treatment and at subsequent points in time. Findings from this study will provide population-based estimates of the proportions of patients suffering various side effects from radiation and surgery.

Three studies examining the impact of prostate-specific antigen testing on recent trends in prostate cancer rates are in press (7-9). Recently, a report to the nation on the status of cancer was published by researchers from the NCI, the American Cancer Society, and the Centers for Disease Control and Prevention using SEER data as well as mortality data from the National Center for Health Statistics (10). Other recent reports deal with how prostate-specific antigen testing has affected...
prostate cancer incidence (11) and changing patterns in the incidence of esophageal and gastric cancer in the United States (12). SEER data have been used in the development of risk-based recommendations for screening women under age 50 (13). Patterns of care studies have been conducted for selected cancers to address issues pertaining to treatment patterns (14). The occurrence of multiple cancers has been studied for selected primary tumors, with one example being testicular cancer (15). Statistical methods have been used in conjunction with SEER data to estimate asymptomatic incidence and duration of prostate cancer (16), and new methods have been developed to evaluate the statistical significance of space-time clusters of cancer (17). Statistical methods have also been developed to estimate lifetime and age-conditional probabilities of developing cancer (18). A number of other studies in which researchers at the various registries have taken a lead role have been performed under the SEER Special Study mechanism. A sampling of the publications stemming from these studies is included in the reference list (19–39).

Other research activities include the development of new statistical methods for cancer surveillance. Methods are being developed to estimate incidence and prevalence rates for the total United States and at the state level using patient survival data from SEER areas and cancer mortality data for the total United States. Methods have also been developed for describing trends in age-adjusted rates and partitioning linear trends (40, 41). Incidence-based mortality methodology has been developed that provides for partitioning cancer mortality trends based on variables identified at diagnosis (42). This methodology also provides a method for assessing the impact of an intervention introduced into the general population on the trend in mortality rates for the cancer expected to be affected. This is only a sampling of recent reports from researchers involved in cancer surveillance at the NCI.

SEER data have been linked to the Medicare database. There is a public use file containing data from the linked file that can be obtained for research purposes. Interested researchers must submit a proposal describing their proposed study. The proposal is reviewed by both the Health Care Finance Administration and the NCI. Details about this file can be obtained on the Web.4

Cancer surveillance at the NCI has been expanded to include the evaluation of patterns and trends in cancer-related risk factors, health behaviors, and health services and their effects on trends in cancer incidence, morbidity, mortality, and patient survival. Many of the studies are being done in SEER areas. Research focuses on: (a) identifying, improving, and developing databases and methods for cancer-related surveillance research; (b) maintaining, updating, and disseminating these databases and methods; and (c) promoting and facilitating their use among investigators within the extramural research community and federal agencies.

Much work has gone into facilitating the use of the SEER database for research purposes. PC-based software has been developed that greatly enhances the ability of those who obtain SEER public use files to analyze the data. A program (SEER*Stat) has been developed that allows for the computation of frequency distributions, incidence rates, trends (percent changes and annual percent changes), survival rates (observed and relative), and tests to assess the statistical significance of trends and differences in rates. This software is included free of charge with the public use files that the SEER Program provides to outside investigators. Non-SEER data can also be loaded into SEER*Stat using a software package called SEER*Prep. SEER*Stat has been adopted with great enthusiasm by the research community. In addition, a Cancer Query System (CANQUES) is accessible at the SEER Web site. It is a large database of 5.7 million precalculated cancer incidence and survival statistics that have been formatted for easy retrieval.

NCI investigators involved in surveillance activities have developed new statistical tools that address important cancer surveillance issues including: (a) a Prevalence System that uses incidence and survival data from 1950 to the present for the State of Connecticut to predict prevalence in the United States; (b) a PC-based system that uses incidence, mortality, and population data to calculate the probability of developing or dying from cancer (DevCan); and (c) a computer modeling program for the PC that is used to estimate the effects of various prevention, screening, and treatment interventions on cancer incidence, prevalence, and mortality (Can*Trol). Further information about these tools is available on the internet.4

NCI staff responsible for managing the SEER Program interact with a number of organizations that are also involved in

---

4 Further information is available at http://www-dccps.ims.nci.nih.gov/arb/.

---

Fig. 2. Percentages and sizes of the populations included in SEER areas.

Fig. 3. Selected characteristics of SEER areas versus the total United States. The populations within the coverage areas are included in the SEER areas. Native Alaskans in Alaska and Arizona Indians are not included.
cancer surveillance or related activities. These include the North American Association of Central Cancer Registries, the American Cancer Society, the American College of Surgeons, the National Cancer Registrars Association, the International Association of Cancer Registries, the International Association for Research on Cancer, and the WHO. We also work with other federal agencies involved in cancer surveillance including the Center for Disease Control and Prevention. These activities include setting standards for data collection by cancer registries, providing for the interchange of ideas and tools for cancer surveillance, facilitating the training and credentialing of tumor registrars, collaborating in the analysis and reporting of cancer rates, and exploring the feasibility of a national cancer reporting system.

Technical assistance is provided to non-SEER registries upon request, subject to available resources. The nature of this assistance includes conducting workshops on data collection, procedures and registry operations and training classes in the use of SEER*Stat. In the latter case, the training includes putting the registry’s database into SEER*Stat, with the remainder of the training devoted to analyzing the registry’s data.

Cancer surveillance activities at the NCI, including those directly involving the SEER Program and those that go beyond the SEER Program, have been reviewed by a committee of researchers from within and outside the institute (43). Recommendations for the future have been made in a number of areas. The first priority is to expand the scope of surveillance research through additional data collection and methods development. Specific activities will include: (a) collection of data on patterns of care, health status, and quality of life as well as cohort studies of newly diagnosed cancer patients for the purpose of documenting levels and trends in these parameters; (b) collection of risk factor and screening data in defined populations, particularly those covered by high-quality cancer registration; (c) development of research methods to measure the dimensions of the cancer burden and factors affecting the burden as well as methods to explain patterns and trends in cancer rates; and (d) exploration of the feasibility and utility of using geographic information systems for geocoding surveillance data and reporting geographic relationships among screening measures, risk factors (including environmental exposures), and improved cancer outcomes.

A second area of focus is to expand the scope of surveillance to improve the representativeness of cancer burden estimates. Specific activities will include expanding SEER population coverage to improve representation of ethnic minority and underserved populations including rural African Americans, Hispanics from Caribbean countries, American Indians, residents of Appalachia and other rural areas, especially those of lower socioeconomic classes; develop methods for improving national estimates of the cancer burden; and work with other organizations involved in cancer surveillance to develop a national cancer surveillance plan.

A third area to be addressed is the production and dissemination of a national report card on the cancer burden. Specific activities will include the collection, analysis, and dissemination of data on important cancer outcomes and trends in risk factors, screening, and treatment to be incorporated into a national cancer report card and the development of improved methods for disseminating information via the report card and other NCI communications.

The fourth area to be addressed is the support of molecular and genetics research for surveillance. Specific activities will include the development of valid tools to assess family history of cancer that will provide for the collection of data on the population prevalence of familial cancers and the investigation of the feasibility of expanding population-based molecular and genetic biomarker studies within the Cancer Surveillance Research Program.

The final area to receive attention is the development of a training strategy for individuals interested in cancer surveillance research. Specifically, training pertaining to the needs of surveillance sciences will be developed along with a plan to incorporate surveillance training as a priority in mechanisms for training cancer prevention and control scientists. Much more detail is provided in the Cancer Surveillance Research Implementation Plan (43), which can be obtained on the Web.8

We have attempted to give a brief overview of cancer surveillance at the NCI with regard to both current activities and future directions. The increase in the scope of surveillance activities including the enhancement of research opportunities will involve the collaborative efforts of federal and state institutions, the academic and medical communities, professional and voluntary health organizations, industry, and patient/public action groups devoted to the fight against cancer.

References


8 The Cancer Surveillance Research Implementation Plan can be obtained at http://cancer.cancer.gov/dci/ip/.
The Surveillance, Epidemiology, and End Results Program: A National Resource

Benjamin F. Hankey, Lynn A. Ries and Brenda K. Edwards


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/8/12/1117

Cited articles
This article cites 34 articles, 9 of which you can access for free at:
http://cebp.aacrjournals.org/content/8/12/1117.full.html#ref-list-1

Citing articles
This article has been cited by 68 HighWire-hosted articles. Access the articles at:
/content/8/12/1117.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.