

# **Cancer Surveillance Research Implementation Plan**

**March 1999**



Prepared by the  
Surveillance Implementation Group (SIG)  
National Cancer Institute  
National Institutes of Health



# Cancer Surveillance Research Implementation Plan

**March 1999**

Prepared by the  
Surveillance Implementation Group (SIG)  
National Cancer Institute  
National Institutes of Health



# TABLE OF CONTENTS

<b>Executive Summary .....</b>	<b>1</b>
<b>Chapter 1: Introduction .....</b>	<b>9</b>
1.1 NCI’s Cancer Surveillance Research Program .....	9
1.2 The Surveillance Implementation Group .....	11
1.2.1 Vision for the Cancer Surveillance Research Program .....	12
1.2.2 Process of the Surveillance Implementation Group .....	13
<b>Chapter 2: NCI’s Cancer Surveillance Research Program: Where Are We Now? .....</b>	<b>15</b>
2.1 The SEER Program .....	15
2.2 Explaining Trends in the Cancer Burden .....	17
2.2.1 Dimensions of the Cancer Burden .....	18
2.2.2 Monitoring Factors that Affect the Cancer Burden .....	20
2.2.3 The Relationship of Risk Factors, Screening, and Treatment to the Cancer Burden .....	22
2.2.4 Why the Cancer Burden is Changing and What Can Be Done to Reduce It .....	24
<b>Chapter 3: The Strategy: Where Do We Go From Here? .....</b>	<b>25</b>
Priority Area 1: Expand the scope of surveillance research through additional data collection and methods development .....	26
Priority Area 2: Expand the scope of surveillance to improve the representativeness of cancer burden estimates .....	31
Priority Area 3: Produce and disseminate a national report card on the cancer burden .....	35
Priority Area 4: Support molecular and genetics research for surveillance .....	37
Priority Area 5: Develop a training strategy for cancer surveillance research .....	38
<b>Appendix A: The NCI Surveillance Implementation Group .....</b>	<b>41</b>
<b>Appendix B: Cancer Surveillance in the United States .....</b>	<b>45</b>
<b>Appendix C: Major Surveys and Databases Used by NCI’s Cancer     Surveillance Research Program .....</b>	<b>51</b>
<b>Appendix D: List of Acronyms .....</b>	<b>59</b>



# EXECUTIVE SUMMARY

**M**onitoring emerging trends in our national cancer burden, and the factors that influence these measures, is extremely important in our efforts to reduce the burden of cancer. The role of the National Cancer Institute's (NCI) Cancer Surveillance Research Program is to study such trends, track the impact of cancer on the general population, and provide information that will enable researchers to generate hypotheses and address questions about observed changes over time. Appropriate decision making in science and in public health depends on reliable information about the effects of our efforts to control cancer. An effective surveillance program drives the cancer control research agenda by identifying opportunities for investment with high payoff in terms of reduced morbidity and mortality.

The Cancer Surveillance Research Program is the most authoritative source of information on cancer incidence, mortality, and stage-specific survival in the United States. Its rigorous quality standards have made it the model for cancer surveillance activities throughout the world. For more than 25 years, NCI's Surveillance, Epidemiology, and End Results (SEER) Program—a major component of the Cancer Surveillance Research Program—has tracked the impact of cancer on the general population, amassing data on more than 2.5 million cancer cases. These data increasingly are being used to answer questions about cancer causation, prevention, treatment, and control. SEER and its predecessor programs have enabled the NCI to elucidate environmental carcinogens, to track the cancer-related effects of tobacco on men and women, to lo-

cate geographic areas with higher than average rates of cancer, to study patterns and outcomes of cancer care, to estimate the cost of cancer, and to identify risk groups for research and public health intervention programs. All of this has been accomplished while maintaining the highest level of confidentiality and privacy. Because of NCI's cancer surveillance program, we know that overall cancer incidence and mortality rates in the United States have begun to decline for the first time in this century, even though rates for certain cancers, such as female lung cancer and non-Hodgkin's lymphoma, are increasing. Opportunities now exist to expand the surveillance program to ensure that it will provide information needed to better characterize the cancer burden, interpret observed changes in trends over time, and enhance researchers' ability to generate hypotheses.

To identify what cancer surveillance research is most needed and how best to advance our knowledge of cancer based on the opportunities available, the Director of the NCI established the Surveillance Implementation Group (SIG), which included 42 leading scientists and experts from within the NCI, other federal agencies, and the extramural community as well as representatives of major NCI review and advisory committees. The SIG was charged with providing advice and recommendations for expanding and enhancing NCI's Cancer Surveillance Research Program. The SIG was asked to identify research directions and priorities and to produce an implementation plan that presented a comprehensive, focused, coherent vision for NCI-funded surveillance research. The framework for the

SIG's deliberations was based on the recent reviews of the cancer control and prevention programs at the NCI, but the group was directed to go beyond these reports to provide scientific guidance about cancer surveillance programs of the future.

The SIG took as its primary directive the recommendations of the Cancer Control Program Review Group (CCPRG). Regarding NCI's surveillance program, the CCPRG recognized the high quality of the data collection, research, and reporting activities and noted that SEER data have been used nationally for many reports on cancer trends and patterns and to facilitate data collection for epidemiologic, cancer control, and genetic studies. To improve the current surveillance program, the CCPRG recommended that the NCI should:

- Expand the SEER Program to include additional populations, more data from patients' medical records and patients themselves, and population data from the SEER regions to monitor individual and societal mediators of cancer.
- Use the SEER expanded data and expertise to produce a timely report card on the cancer burden.

The members of the SIG agreed that the implementation plan should extend beyond the SEER Program expansion recommended by the CCPRG. They also concurred that NCI's surveillance program should provide answers to important questions about the national cancer burden. Addressing these questions will require substantial enhancements to the current NCI surveillance system.

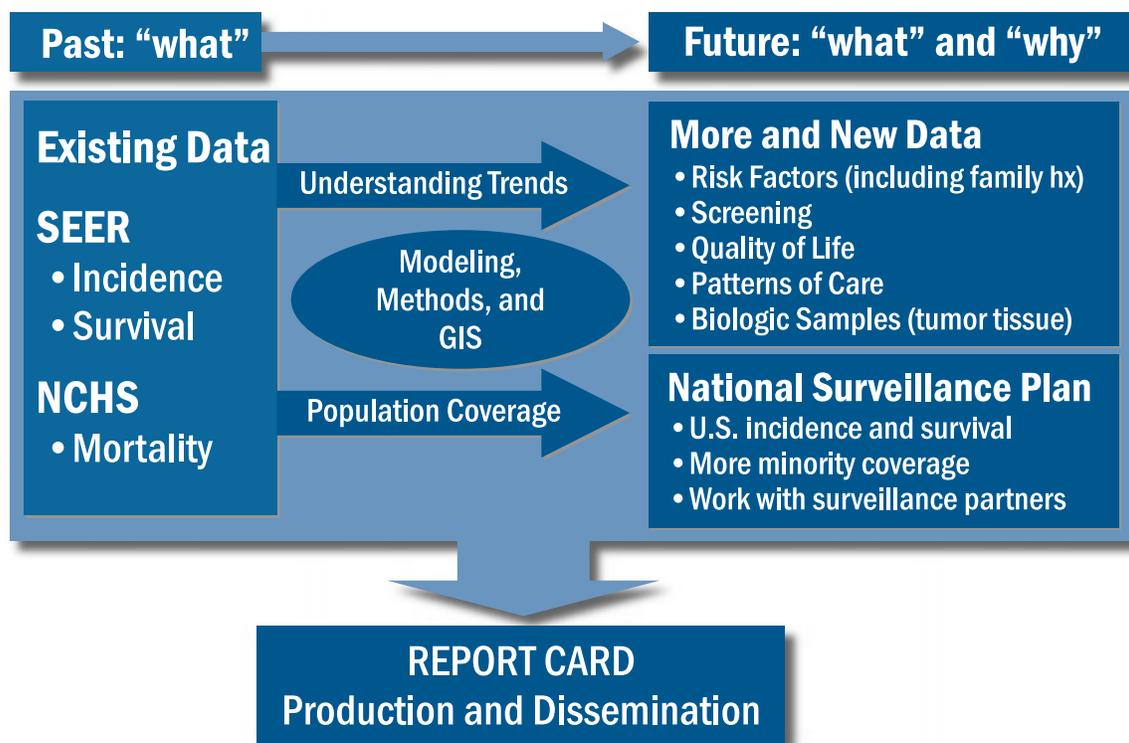
Clearly, the SEER Program—a continuing model of excellence in cancer surveillance throughout the world—will remain the core of the expanded system. As new tools are developed, SEER will be connected to data collection mechanisms that probe deeply into the causes of cancer rates and trends with a consistent focus on defined populations as the point of reference. These defined population studies in the SEER areas (or other areas of high-quality registration) will collect data on prevention, risk factors, screening, and treatment interventions. This will involve data collection on cohorts of patients over time, and will include measures of health status, patterns of care, and quality of life. These

### Some Key Questions to be Addressed by an Expanded Surveillance System

- What is the cancer burden on the population?
- How is that burden changing over time, in different areas of the country, and in different racial/ethnic groups?
- What are the factors that influence both the rates and trends?
- How do these factors vary over time, geographically, and among racial/ethnic groups?
- Are the burden and its trends uniform over the population?
- Why did the rates and trends change?

data will provide information for understanding specific questions concerning cancer rates, such as the impact of early detection on colorectal cancer mortality rates. The data collection models, similar to NCI's Breast Cancer Surveillance Consortium and Prostate Cancer Outcomes Study, will require increasing connections of the surveillance program with the extramural epidemiology and treatment programs at the NCI. As other registries around the country match the SEER standard, data can be pooled. Any expansion of SEER activities will be taken in concert with a newly developed long-term surveillance research plan that includes collaborations with other organizations involved in cancer surveillance. Finally, the strong research structure serving as the underpinning of SEER will continue, with major methodologic efforts in modeling rates and trends along with new efforts in geographic information systems, development of improved approaches to generating national estimates of the cancer burden, and new research on familial and genetic components of cancer surveillance.

The SIG's vision for an expanded surveillance program can be illustrated with the help of the framework on the following page that indicates where cancer surveillance has been and where it can go as the enhancements recommended by the SIG are implemented.



To accomplish its charge, the SIG met in April, August, September, November, and December 1998, and smaller working groups participated in numerous conference calls during this 9-month period. After developing a vision for the Cancer Surveillance Research Program, the SIG reviewed NCI's portfolio of surveillance research, the balance of the current research across topic areas, and the distribution of the program's budget across projects. The SIG then determined which of the CCPRG's recommendations already had been partially or completely implemented, identified emerging issues from research in progress, and generated a list of recommendations. Through a consensus-building process, the SIG narrowed this list to 12 cancer surveillance research opportunities, which were organized within five overarching priority areas. These research opportunities range from expansion of data collection to development of tools for analyzing surveillance data to establishment of linkages among cancer and other health-related databases. These recommendations provided the foundation for development of this Implementation Plan.

The SIG emphasized that formation of strategic partnerships in the implementation of this research agenda will be critical to its success. The NCI must collaborate with partners in both the public and private sectors, such as the Centers for Disease Control and Prevention (including the National Center for Health Statistics and the National Center for Chronic Disease Prevention and Health Promotion), the Health Care Financing Administration, the American Cancer Society, the Commission on Cancer of the American College of Surgeons, the North American Association of Central Cancer Registries, the National Coordinating Council for Cancer Surveillance, grantees, and others.

The Executive Summary presents an overview of the SIG's recommendations and the 12 research opportunities. The full report provides more background on the cancer surveillance research program and the in-depth rationale for the recommendations. The cancer surveillance research opportunities identified by the SIG are summarized on the following pages within the five priority areas.

## Priority Area 1:

### Expand the scope of surveillance research through additional data collection and methods development.

#### Research Opportunities

1. **Support the collection of data on patterns of care, health status, morbidity, and quality of life as well as cohort studies of newly diagnosed registered cancer patients for the purpose of documenting levels and trends in these parameters.** *(The cost for this effort is expected to be high; work should be initiated within the next 1-2 years.)*

To more fully assess the Nation's cancer burden, data on patterns of care and cancer outcomes beyond incidence, survival, and mortality are needed. To this end, the SIG recommends research to determine the best measures of patterns of care, health status, morbidity, and quality of life. Support is needed for the collection of these measures in incidence cohorts within established population-based cancer registries. Such research should follow the model of the Prostate Cancer Outcomes Study.

2. **Support the collection of risk factor and screening data in defined populations, particularly those covered by high-quality cancer registration.** *(The cost of this effort is expected to be high; work should be initiated within the next 1-2 years.)*

To explain trends in dimensions of the cancer burden in terms of lifestyle and behaviors such as smoking, diet, physical activity, and participation in screening, representative data about risk factors and screening must be collected from defined populations where SEER-quality registration occurs. The SIG recommends that the surveillance program initiate studies to link risk factor and screening data directly to cancer outcome data. This could be accomplished through the expansion of existing national health surveillance systems, such as the Current Population Survey, the National Health Interview Survey, and the Behavioral Risk Factor Surveillance System. The NCI could support the development of selected health surveillance systems and databases with geographically specific data as well as efforts to collect risk

factor and screening data in defined local populations linked to SEER-quality cancer registration systems. Research on the relationship of screening practices to cancer outcomes could be implemented by developing consortia modeled after the Breast Cancer Surveillance Consortium. Specifically, the SIG supports the development of a Colorectal Cancer Surveillance Consortium as a high priority.

3. **Develop research methods to measure dimensions of the cancer burden and factors affecting the burden as well as methods to explain patterns and trends in cancer rates.** *(The cost for this effort is expected to be moderate; work should be initiated within the next 1-2 years.)*

As data on additional measures of the cancer burden and related factors are collected, appropriate tools for managing and analyzing these data will be needed. The SIG recommends that research be conducted to improve methods for measuring quality of life, quality of care, health status, morbidity, family history, environmental exposures, behaviors (smoking, diet, physical activity, tobacco use), screening, treatment, and cancer biology as well as methods and models for relating variables and predicting outcomes. In addition, statistical modeling and methods are needed to solve quantitative problems in cancer surveillance and to study the impact of cancer control interventions on the cancer burden. The Cancer Intervention Surveillance NETwork Modeling (CISNET) concept is a good example of this latter type of research.

4. **Explore the feasibility and utility of employing geographic information systems for geocoding surveillance data and reporting geographic relationships among screening measures, risk factors (including environmental exposures), and improved cancer outcomes. Methods need to be developed for assuring data confidentiality.** *(The cost for this effort is expected to be moderate; work should be initiated within the next 1-2 years.)*

Research is needed on the utility of geographic information systems (GIS) as an innovative addition to the cancer surveillance infrastructure. By capturing county, census tract, and the coordinates (latitude and longitude) of where cancer patients live in the SEER system, the number of linkages with other geographically based data systems becomes maximized. Furthermore, such methods potentially allow linkages for the entire country for incident and fatal cancers and can thus be used to help explain local phenomena and trends. Issues of confidentiality must be resolved to extend the utility of this technology.

To support cancer surveillance efforts, there is a need for geocoded data on socioeconomic status, residential neighborhood characteristics, and environmental factors. Validation of geocoding also is needed to facilitate quantification of the relationships among risk factors, screening, and cancer outcomes for defined geographic areas. The SIG recommends that the Cancer Surveillance Research Program improve capacity for GIS and cancer surveillance in SEER and other NCI data systems. In addition, the SIG recommends the support of workshops on GIS and cancer surveillance.

## Priority Area 2:

### Expand the scope of surveillance to improve the representativeness of cancer burden estimates.

#### Research Opportunities

5. **Expand NCI's surveillance program to improve representation of ethnic minority and underserved populations.** *(The cost of this effort is expected to be high; work should be initiated within the next 1-2 years.)*

The SEER Program has added new populations over time to include under-represented segments of the population—particularly African Americans, Hispanics (Mexican Americans), Asians, and Pacific Islanders. The SEER Program also provides technical assistance to newer registries among American Indians and the established Registry for Alaska Natives. Currently, SEER is not a fully representative sample of rural African Americans, Hispanics from Caribbean countries, American Indian populations, and residents of Appalachia and other rural areas, especially those of lower socioeconomic classes. The recommended expansion should be accomplished by working with the Centers for Disease Control and Prevention's National Program of Cancer Registries, the North American Association of Central Cancer Registries, and states to collect and report SEER-quality incidence, mortality, and survival data on one or more of these populations.

6. **Explore methods for developing improved national estimates of the cancer burden.** *(The cost for this effort is expected to be low; work should be initiated this year.)*

The 14 percent of the U.S. population covered by SEER registries does not represent completely the national cancer burden in a statistical sense. Short of an eventual goal of complete cancer registration in the United States, research is needed to develop and consolidate the following two complementary approaches to improving national estimates of the cancer burden: (1) sampling methods and estimation techniques that, with appropriate data, would generate a probabilistically based estimate of the cancer burden; and (2) techniques for modeling national rates based on SEER and other population-based registries.

7. **Work with partners to develop a National Cancer Surveillance Plan.** *(The cost of this effort is expected to be low; work should be initiated this year.)*

There currently are several national efforts in cancer surveillance, each of which was designed for a specific purpose. The goal of a comprehensive National Cancer Surveillance Plan can be achieved only by an effective working partnership among the agencies and organizations responsible for these surveillance efforts and their primary end users. The SIG recommends that NCI's Cancer Surveillance Research Program work with these partners in the context of the National Coordinating Council for Cancer Surveillance to forge a successful alliance to develop and implement a plan to achieve complete and comprehensive cancer surveillance for the country.

## Priority Area 3: Produce and disseminate a national report card on the cancer burden.

### Research Opportunities

8. **Collect, analyze, and disseminate data on important cancer outcomes and trends in risk factor and screening behaviors as well as explanations for these trends in a National Cancer Report Card.** *(The cost for this effort is expected to be low; work should be initiated this year.)*

A National Cancer Report Card is needed to evaluate the Nation's performance in its ongoing efforts to confront the substantial burden of cancer on the U.S. population. The Report Card needs to include traditional cancer statistics on the burden from major cancers and all cancers together as well as trends in these rates. The Report Card also needs to capture the results of analytic research produced by the expanded Cancer Surveillance Research Program as well as research results from other NCI programs, other federal agencies, and the extramural community that explain variations in the cancer burden and trends in the factors affecting it.

The Report Card needs to be presented in a format that is accessible to the general public, scientists, legislators, and policymakers.

9. **Develop a strategy for improved dissemination of information on the cancer burden via the Report Card and other NCI communications.** *(The cost of this effort is expected to be low to moderate; work should be initiated this year.)*

Better methods to disseminate the findings of the Cancer Surveillance Research Program through other NCI communications beyond the Report Card are needed. These methods should ensure that more detailed data are made available to NCI and extramural investigators to facilitate the generation of hypotheses for epidemiologic studies and behavioral interventions. The data also should be available for studies to better explain the nature of the cancer burden and trends over time. The SIG recommends the development of an information dissemination strategy for the surveillance program.

## Priority Area 4: Support molecular and genetics research for surveillance.

### Research Opportunities

10. **Develop valid tools to assess family history of cancer and collection of data on the population prevalence of familial cancers.** *(The cost of this effort is expected to be moderate; work should be initiated this year.)*

Tracking the population-based prevalence of family history of cancer, including known major inherited cancer syndromes, will provide critical data for both cancer research and public health planning. There is a need to develop standardized, validated instruments for assessing family history of cancer as part of the surveillance program as well as procedures for assuring the confidentiality of the individual-level data that are collected. The impact of family history and cancer susceptibility genes on population-based cancer outcomes also should be examined.

11. **Investigate the feasibility of expanding population-based molecular and genetic biomarker studies within the Cancer Surveillance Research Program.** *(The cost for this effort is expected to be moderate to high; work should be initiated within the next 1-2 years.)*

The surveillance of cancer-related genetic and molecular biomarkers would be strengthened by developing criteria for selecting specific biomarkers, enhancing infrastructure for collecting and archiving DNA-containing biospecimens, and addressing the ethical, legal, and social issues associated with these activities. The SIG recommends that the NCI initiate methodologic and pilot studies to evaluate the feasibility of biomarker studies within the Cancer Surveillance Research Program. Also, the current infrastructures for rapid case ascertainment and data tracking systems should be enhanced.

## Priority Area 5: Develop a training strategy for cancer surveillance research.

### Research Opportunity

**12. Identify specific training needs related to surveillance sciences and develop a plan to incorporate surveillance into mechanisms for training cancer prevention and control scientists.** *(The cost for this effort is expected to be moderate; work should be initiated within the next 1-3 years.)*

The SIG noted that no formal training in surveillance sciences currently exists. With the growing demand for collecting additional data

and the increasing complexity of analyzing such data, there is a need for training programs in surveillance sciences. Specific training needs related to cancer surveillance should be identified. The Cancer Surveillance Research Program should play a significant role in working with members of the National Coordinating Council on Cancer Surveillance, professional organizations, and NCI's training and fellowship programs to develop and implement a plan to address these training needs.

The relationship between the 12 research opportunities and the recommendations of the Cancer Control Program Review Group is depicted in the table below, along with a brief description of the initiatives and activities proposed by the SIG for implementing these research opportunities.

Recommendations of CCPRG (9/97)	Recommendations of SIG (2/99)	Implementation Initiatives of SIG (2/99)	Timetable to Initiation
I. Expand the Cancer Surveillance Research Program (SEER)			
A. Include more data on cancer patients and populations	1. Collect data on patterns of care, quality of life, and morbidity in cohorts of registered cancer patients.	1. Support site-specific cancer cohorts.	1-2 years
	2. Collect data on risk factors and screening in defined populations.	2. Innovative methods to connect risk factor and screening data to cancer outcome data in defined registered populations. Colorectal Cancer Surveillance Consortium. Collaborate with existing national and/or regional health surveillance systems to collect data and improve data linkage.	1-2 years
	3. Develop research methods to improve measurement of cancer burden and trends.	3. RFA for Cancer Intervention and Surveillance Modeling NETwork (CISNET). Methodologic studies to improve ascertainment of risk factors (e.g., SES, tobacco, diet) and develop additional measures of cancer burden.	1-2 years
	4. Explore geographic information system (GIS) methods.	4. Improve capacity for GIS in SEER and other Cancer Surveillance Research Program data systems. Support workshops on GIS and cancer surveillance.	1-2 years

Recommendations of CCRPG (9/97)	Recommendations of SIG (2/99)	Implementation Initiatives of SIG (2/99)	Timetable to Initiation
B. Include additional populations	5. Expand to improve representation of ethnic minority and underserved populations.	5. Supplementary funding for selected non-SEER regions/states through collaborative initiatives with partner members of the National Coordinating Council for Cancer Surveillance. Support methods and feasibility studies of expansion to improve national cancer data for selected populations.	1-2 years
	6. Explore methods for improved national estimates of cancer burden.	6. Form working group of experts from within the NCI and extramural community to investigate concepts and fund small technical studies when needed. Hold workshop on statistical designs for probabilistic estimates of the national cancer burden.	This year
	7. Work with partners on National Cancer Surveillance Plan.	7. Working meetings and workshops with members of the National Coordinating Council for Cancer Surveillance.	This year
II. Use Expanded Data to Produce a Timely Report Card on the Cancer Burden	8. Proceed with production of Report Card for Year 2000.	8. In-house budget to design, publish, and distribute report.	This year
	9. Develop plan to improve dissemination of surveillance data.	9. In-house budget.	This year
	10. Develop tools to assess family history and prevalence of familial cancers.	10. Validate family history of cancer questionnaire through surveillance survey contract.  Conduct prevalence study of family history of cancer, if feasible, through surveillance survey contract.	This year  2-3 years
	11. Explore the feasibility of collecting bio-specimens to support population-based molecular and genetic research.	11. SEER Special Studies to explore feasibility of collecting biospecimens. Ancillary exploratory studies conducted as integrated initiatives with NCI grant programs.	1-2 years
	12. Identify training needs and develop plan to incorporate surveillance sciences into training mechanisms.	12. In-house activity with partner members of the National Coordinating Council for Cancer Surveillance, professional organizations, and NCI's training and fellowship programs.	1-3 years

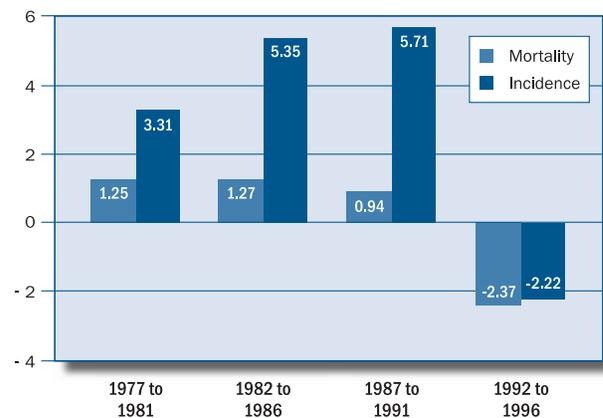
# 1 Introduction

The ultimate goal of research sponsored by the National Cancer Institute (NCI) is to reduce the burden exacted by cancer by preventing or curing it. Basic laboratory and epidemiologic research are elucidating the causes, prevention, and cures of cancer. Intervention research is demonstrating how the cancer burden can be affected by manipulating risk factors and health-related behaviors known to cause, prevent, or cure cancer. Surveillance research is describing how cancer rates respond to interventions known to be effective. Control of cancer requires such discoveries as well as translation of these discoveries into effective interventions and delivery of these interventions to the population. It also requires synthesis of knowledge in order that informed decisions can be made regarding the cancer research agenda and public health policy.

Genes involved in the progression of the major cancers have been identified and their translation into effective interventions will be important in controlling cancer. During the last two decades, screening has been introduced for three of the most common cancers: breast, colon, and prostate. In fact, the observed increases in incidence rates in the 1980s for breast and prostate cancer may be attributable in part to increases in screening. In addition, new therapies for preventing and treating certain cancers have been introduced, such as taxol, tamoxifen, and herceptin. At least one of these—tamoxifen—has been shown to have the potential to prevent breast cancer. Perhaps most importantly, education of the public regarding the relationship between cancer and lifestyle has had a positive impact on cancer-related behaviors, such as a reduction in smoking rates. These major breakthroughs in our understanding of cancer, and the resulting growth in our arsenal of weapons against cancer, are beginning to

be reflected in declines in cancer incidence. Already, we are observing declines in the mortality rate for many cancers, including cancers of the breast, prostate, male lung, and colorectum. The annual age-adjusted cancer death rate of the U.S. population fell between 1992 and 1996, the first sustained decline since national record keeping was instituted in the 1930s.

**Figure 1.**  
Percent Change in U.S. Cancer Mortality and SEER Cancer Incidence Rates for Each 5-Year Period, 1977-1996

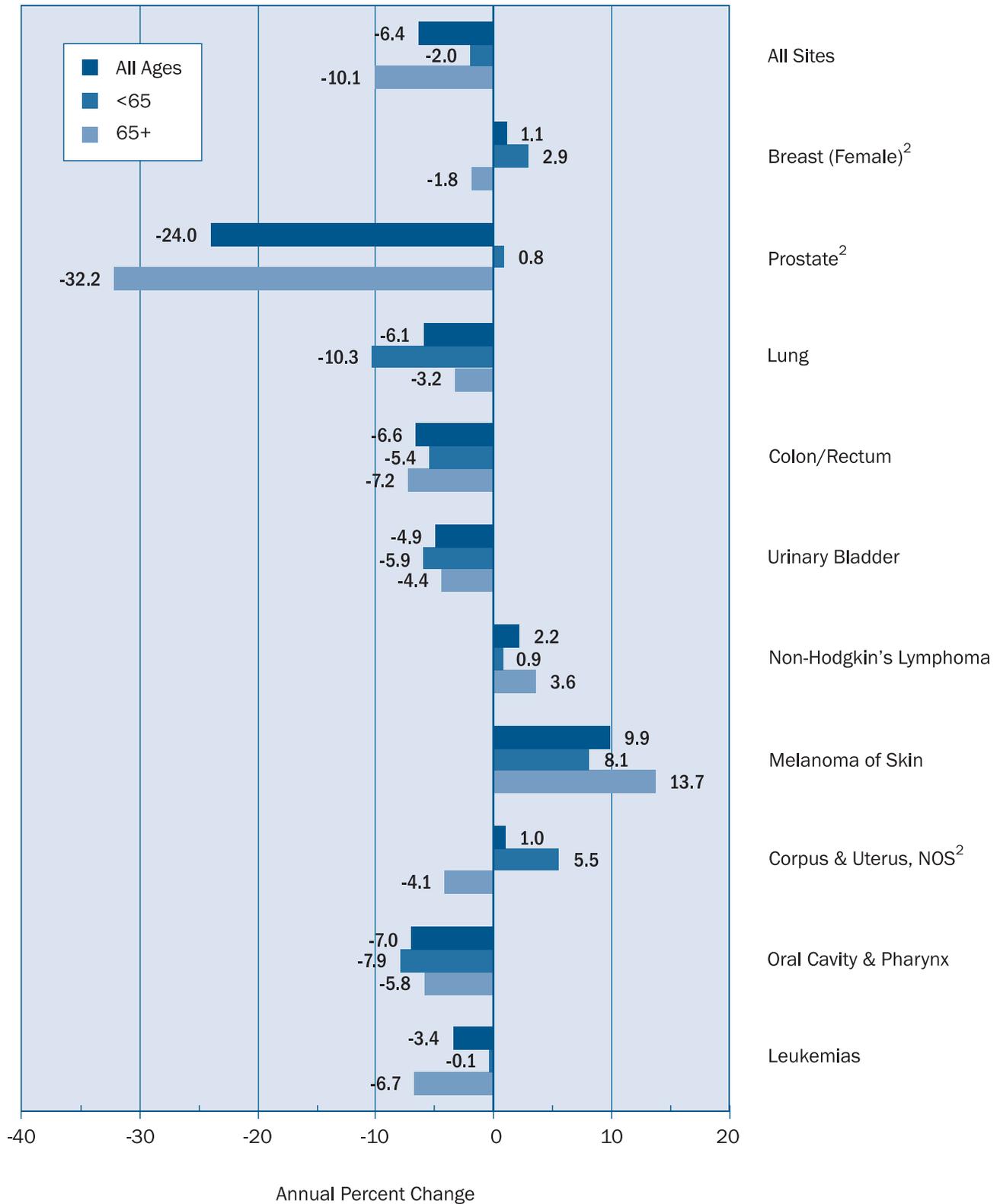


Source: SEER database, 1999

## 1.1 NCI's Cancer Surveillance Research Program

The role of NCI's Cancer Surveillance Research Program (CSRP) is to study cancer rates and trends, track the impact of cancer on the general population, and provide information that will generate hypotheses and address questions about changes in trends over time. Appropriate decision making in science and in public health depends on reliable

**Figure 2.**  
**SEER Cancer Incidence<sup>1</sup>, Percent Change 1992-1996**  
**Trends for Top 10 Sites, All Ages, All Races**



<sup>1</sup> Incidence rates are age-adjusted to the 1970 U.S. standard million population.

<sup>2</sup> Percent change is based on gender-specific rates.

information about the effects of our efforts to control cancer. An effective surveillance program drives the cancer control research agenda by identifying opportunities for investment with high payoff in terms of reduced morbidity and mortality.

The CSRPs is the most authoritative source of information on cancer incidence, mortality, and stage-specific survival in the United States. Its rigorous quality standards have made it the model for cancer surveillance activities throughout the world. Since 1973, the CSRPs's Surveillance, Epidemiology, and End Results (SEER) Program has tracked the impact of cancer on the general population, amassing data on more than 2.5 million cancer cases. These data increasingly are being used to answer questions about cancer causation, prevention, treatment, and control. For more than 25 years, SEER and its predecessor programs have enabled the NCI to elucidate environmental carcinogens, to track the cancer-related effects of tobacco on men and women, to locate geographic areas with higher than average rates of cancer, to study patterns and outcomes of cancer care, to estimate the cost of cancer, and to identify risk groups for research and public health intervention programs. All of this has been accomplished while maintaining the highest level of confidentiality and privacy.

The CSRPs has done an outstanding job of characterizing the cancer burden and identifying trends in incidence, mortality, and survival rates for specific cancer sites. Because of NCI's cancer surveillance program, we know that overall cancer incidence and mortality rates in the United States have begun to decline for the first time in this century, even though rates for certain cancers (e.g., female lung cancer and non-Hodgkin's lymphoma) are rising.

Opportunities now exist to expand the surveillance program to ensure that it will provide information needed to better characterize the cancer burden, interpret observed changes in trends over time, and enhance researchers' ability to generate hypotheses. Recent advances in information technology provide opportunities to link cancer surveillance data with different types of health-related information on populations, which can provide a powerful tool for analyzing factors that influence cancer rates (risk factors, screening, treatment, and health practices) and for planning and evaluating population-based prevention and control interventions.

## 1.2 The Surveillance Implementation Group

To identify what cancer surveillance research is most needed and how best to advance our knowledge of cancer based on the opportunities available, the Director of the NCI established the Surveillance Implementation Group (SIG), which included 42 leading scientists and experts (see Appendix A for a list of the SIG members) from within the NCI, other federal agencies, and the extramural community as well as representatives of major NCI review and advisory committees. The SIG was charged with providing advice and recommendations for expanding and enhancing NCI's Cancer Surveillance Research Program. The SIG was asked to identify research directions and priorities and to produce an implementation plan that was national in scope and that presented a comprehensive, focused, coherent vision for NCI-funded surveillance research. The framework for the SIG's deliberations was based on the recent reviews of the cancer control and prevention programs at the NCI, but the group was directed to go beyond these reports to provide scientific guidance about cancer surveillance programs of the future.

The SIG took as its primary directive the recommendations of the Cancer Control Program Review Group (CCPRG), convened in December 1996 by the Director of the NCI and the NCI Board of Scientific Advisors (BSA) to evaluate the full scope of current and past activities of the Institute's cancer control research program. Regarding NCI's surveillance program, the CCPRG recognized the high quality of the data collection, research, and reporting activities and noted that SEER data have been used nationally for many reports on cancer trends and patterns and to facilitate data collection for epidemiologic, cancer control, and genetic studies. To improve the current surveillance program, the CCPRG recommended that the NCI should:

- Expand the SEER Program to include additional populations, more data from patients' medical records and patients themselves, and population data from the SEER regions to monitor individual and societal mediators of cancer.

- Use the SEER expanded data and expertise to produce a timely report card on the cancer burden.

### 1.2.1 Vision for the Cancer Surveillance Research Program

The SIG members interpreted the recommendations of the CCPRG in light of their **vision** for the CSRP. The members agreed that the implementation plan should extend beyond the SEER Program expansion recommended by the program review group. They also concurred that addressing the many questions relating to the national cancer burden will require substantial changes within the current NCI surveillance system. Clearly, the SEER Program—a continuing model of excellence in cancer surveillance throughout the world—will remain the core of the expanded system. As new tools are developed, the SEER system will be connected to data collection mechanisms that probe deeply into the causes of cancer rates and trends with a consistent focus on defined populations as the point of reference.

These defined population studies in the SEER areas (or other areas of high-quality registration) will collect data on prevention, risk factors, screening, and treatment interventions. This will involve data collection on cohorts of patients over time, and will include information on health status, patterns of care, and quality of life. These data will provide information for understanding specific questions concerning cancer rates, such as: Why are colorectal cancer mortality rates decreasing? What impact does early detection have on colorectal cancer mortality rates? These data collection models, similar to NCI's Breast Cancer Surveillance Consortium and Prostate Cancer Outcomes Study, will require increasing connections of the surveillance program with the epidemiology and treatment programs at the NCI. As other registries around the country match the SEER standard, data can be pooled. Any expansion of SEER activities will be taken in concert with a newly developed long-term surveillance research plan that includes collaborations with other organizations involved in cancer surveillance. Finally, the strong research structure serving as the underpinning of SEER will continue, with major method-

### Vision Statement for NCI's Cancer Surveillance Research Program

The National Cancer Institute's Cancer Surveillance Research Program (CSRP) is the internationally recognized standard of excellence for the comprehensive surveillance of cancer. It characterizes the cancer burden borne in the United States over time by integrating traditional statistics on persons with cancer with the widest possible collection of cancer-related data in the general population. Based on a sound foundation of applied and methodologic research, the CSRP measures progress in reducing the Nation's cancer burden. It simultaneously provides a stimulus for etiologic research and identifies opportunities for intervention research and public health applications.

ologic efforts in modeling rates and trends along with new efforts in geographic information systems, development of improved approaches to generating national estimates of cancer burden, and new research on familial and genetic components of cancer surveillance.

The SIG emphasized that in pursuing the vision for the surveillance program and the goal of complete cancer registration, the NCI and its CSRP must collaborate with partners in national cancer surveillance such as the Centers for Disease Control and Prevention (CDC), including the National Center for Health Statistics (NCHS) and the National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP); the Health Care Financing Administration (HCFA); the American Cancer Society (ACS); the Commission on Cancer of the American College of Surgeons (ACoS); the North American Association of Central Cancer Registries (NAACCR); the National Coordinating Council for Cancer Surveillance (NCCCS); extramural grantees; and others. The involvement of these partners is essential because each organization has its own purpose and niche in the overall national cancer surveillance scheme. The SIG also recommended that the NCI provide feedback to the Institute scientists, extramural researchers, and health policy administrators to improve their understanding of how interventions and public health applications of research results are affecting the national cancer burden.

## 1.2.2 Process of the Surveillance Implementation Group

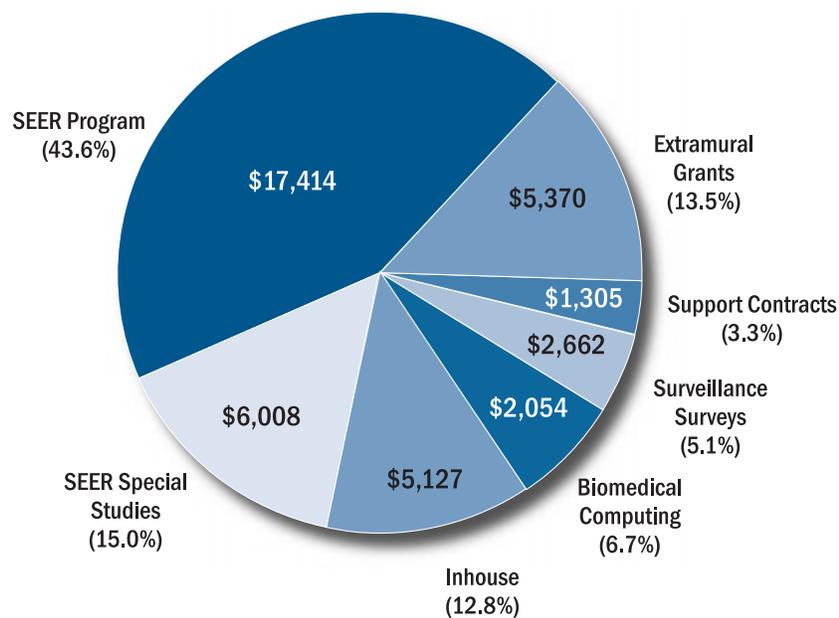
To accomplish its charge, the full SIG met five times during 1998—April, August, September, November, and December—as well as by conference calls for smaller working groups. After establishing consensus on the vision for the CSRP, the SIG analyzed NCI's portfolio of cancer surveillance research, the balance of current research across topic areas, and the distribution of funding across projects and funding mechanism. The NCI spent \$39.9 million on cancer surveillance research in Fiscal Year 1998, which represents approximately 1 percent of NCI's total budget. About half of the CSRP budget was allocated to the SEER Program.

The SIG then determined which of the CCPRG's recommendations already had been partially or completely implemented, identified emerging issues

from research in progress, and generated a list of recommendations. Through a consensus-building process, the SIG members modified and refined this list until it contained 12 cancer surveillance research opportunities within five priority areas. The research opportunities range from expansion of data collection to development of analysis tools to establishment of linkages among cancer and other health-related databases. These recommendations provided the foundation for development of this Cancer Surveillance Research Implementation Plan.

The next chapter of this report provides an overview of NCI's current surveillance research program. Chapter 3 describes the SIG's proposed strategy for implementing the 12 research opportunities for enhancing the surveillance program. A brief rationale for each of the research opportunities is provided in Chapter 3, along with an estimate of the approximate level of investment required and a time frame for initiating the research.

**Figure 3.**  
Fiscal Year 1998 Budget Allocation for NCI's Cancer Surveillance Research Program



**Total FY98 CSRP Budget = \$39,940,000**



# 2 NCI's Cancer Surveillance Research Program: Where Are We Now?

## 2.1 The SEER Program

The SEER Program is a primary component of NCI's CSR, constituting approximately half of its budget. SEER collects and publishes cancer incidence and survival data from population-based cancer registries covering about 14 percent of the U.S. population. The SEER database contains information on more than 2.5 million persons with *in situ* and invasive cancers diagnosed between 1973 and 1996. The quality of the program provides a model for cancer registries throughout the world.

Currently, 25 percent of the American Hispanic population, 41 percent of the Asian/Pacific Islanders population (43 percent of all Chinese Americans and 60 percent of all Japanese Americans), 27 percent of the American Indian population including Alaska Natives living in Alaska, and 12 percent of the African American population reside in SEER areas. The SEER Program is reasonably representative of the U.S. population for purposes of national cancer surveillance based on analyses of mortality trends in SEER areas compared to the total United States. However, certain minority and medically underserved populations are not fully represented in SEER.

The SEER Program is an important population-based laboratory for epidemiologic and surveillance research carried out at the NCI and across the

United States, and it is a valuable resource for researchers, clinicians, public health officials, community groups, and others. In addition to collecting and reporting cancer statistics, the SEER Program monitors 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 provides continuing information on changes over time in the extent of disease at diagnosis, trends in therapy, and associated changes in patient survival. It promotes studies designed to identify factors amenable to cancer control interventions, such as: (1) environmental, occupational, socioeco-

**Figure 4.** Proportion of U.S. Racial/Ethnic Groups in SEER and Non-SEER Registries

Race/Ethnicity	SEER Registries	Non-SEER Registries	Total CINA* (19 Poolable Registries)
White	12.5%	24.7%	37.2%
Black	12.2%	19.1%	31.3%
American Indian, Eskimo, and Aleut	26.6%	10.6%	37.2%
Chinese	43.5%	17.7%	61.2%
Filipino	49.1%	29.4%	78.5%
Japanese	59.8%	20.3%	80.1%
Vietnamese	31.2%	31.6%	62.8%
Korean	33.7%	24.9%	58.6%
Other Race	28.2%	28.7%	56.9%
All Races	14.0%	24.1%	38.1%
Hispanic, All Races	25.0%	26.3%	51.3%

\* *Cancer in North America, 1990-1994* is a 1998 publication of the North American Association of Central Cancer Registries (NAACCR). Only data deemed to be of very good quality are pooled to compute the combined cancer incidence rates for the United States. The specific criteria are in CINA 1990-1994 Volume One: Incidence, NAACCR 1998 (<http://www.naacr.org>).

## The Surveillance, Epidemiology, and End Results (SEER) Program

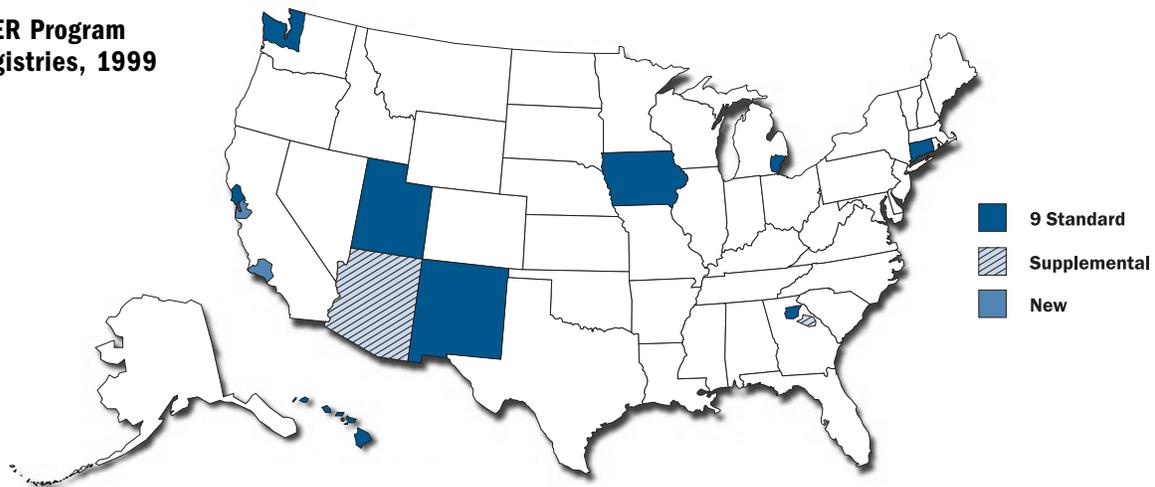
The SEER Program is the most authoritative source of information on cancer incidence and survival in the United States. Case ascertainment for SEER began on January 1, 1973, in the states of Connecticut, Iowa, New Mexico, Utah, and Hawaii and the metropolitan areas of Detroit and San Francisco-Oakland. In 1974-1975, the metropolitan area of Atlanta and the 13-county Seattle-Puget Sound area were added. In 1978, 10 predominantly black rural counties in Georgia were added, followed in 1980 by the addition of American Indians residing in Arizona. Three additional geographic areas participated in the SEER Program prior to 1990: New Orleans, Louisiana (1974-1977); four counties in New Jersey (1979-1989); and Puerto Rico (1973-1989). The National Cancer Institute also began funding a cancer registry that, with technical assistance from SEER, collects information on cancer cases among Alaska Native populations residing in Alaska. In 1992, the SEER Program was expanded to increase coverage of minority populations, especially Hispanics, by adding Los Angeles County and four counties in the San Jose-Monterey area south of San Francisco. The SEER Program currently collects and publishes cancer incidence and survival data from 11 population-based cancer registries and 2 supplemental registries covering approximately 14 percent of the U.S. population. Information on more than 2.5 million cancer cases is included in the SEER database, and approximately 160,000 new cases are accessioned each year within the SEER catchment areas. The

SEER registries routinely collect data on patient demographics, primary tumor site, morphology, stage at diagnosis, first course of treatment, and followup for vital status. The mortality data reported by SEER are provided by the National Center for Health Statistics.

The SEER Program is considered the gold standard for cancer registries around the world. Quality control has been an integral part of SEER since its inception. Every year, quality control site visits are conducted in each SEER area to evaluate the consistency of the data and to determine if all the resident cases are being reported (SEER's standard for case ascertainment is 99 percent). A sample of the cases are reabstracted to evaluate the validity of each item to ensure that the data were abstracted from the hospital medical record accurately. Computer edits also are used by the registries to ensure accurate and consistent data.

Updated annually and provided as a public service in print and electronic formats, SEER data are used by thousands of researchers, clinicians, public health officials, legislators, policymakers, community groups, and the public. SEER data are used to identify geographic and population differences in cancer patterns and to study possible links between cancer incidence and occupations, the environment, and lifestyle influences. The database is being used to study the quality of life of cancer patients as well as outcomes of cancer care. SEER data also are used to assess the extent to which state-of-the-art treatments are being used across the country.

**SEER Program Registries, 1999**



nomic, dietary, and health-related exposures; (2) screening practices, early detection, and treatment; and (3) determinants of the length and quality of patient survival. The SEER registries conduct a number of these studies through their Special Studies effort. Over the past 12 years, SEER Special Studies have been used as a timely and efficient mechanism to answer important questions, providing more in-depth data to describe and help explain trends in cancer, answer specific research questions, and improve the quality and operational aspects of population-based registries. A specific advantage of these studies is that they make use of already existing population-based SEER registries and the expertise of their investigators and managers.

The SEER Program is the only comprehensive source of population-based information in the United States that includes stage of cancer at the time of diagnosis, and survival rates within each stage. The SEER Program disseminates cancer incidence, mortality, and survival data through an annual cancer statistics review, monographs on relevant topics, the SEER Web Site at <http://www-seer.ims.nci.nih.gov/>, various software packages (e.g., SEER\*Stat, SEER\*Prep), and a public-use datafile of more than 2.5 million cancer records from which individual patient identifiers have been removed. SEER data and resources are provided free of charge to extramural researchers.

Most states depend on the SEER database when preparing reports of stage-specific survival. The ACS also relies on the data collected by the SEER Program to make its annual and widely quoted estimates of the coming year's cancer incidence and mortality rates. In addition, the NAACCR uses SEER data as its standard to evaluate completeness in state cancer registries.

## 2.2 Explaining Trends in the Cancer Burden

The data collected by SEER are critical to our understanding of the cancer burden. Much of the information needed to explain the trends that emerge from SEER comes from research within CSRPs Applied Research Branch (ARB) and the extramural research community. Capacity to explain the under-

### SEER Special Studies—A Key to Cancer Surveillance Research

Originally implemented to answer the need for in-depth data beyond that routinely collected on cancer cases, the SEER Special Studies mechanism has been tailored to answer specific research questions, such as comparisons of treatment outcomes or risk factors. The studies are developed by extramural investigators associated with the SEER Program who work collaboratively with NCI scientists to address specific surveillance issues. With a time frame of 1-2 years, these studies can provide a rapid response to questions of national importance. For example, during a controversial period of the Primary Prevention Trial of Breast Cancer, a SEER Special Study was used to obtain improved estimates of the risks associated with tamoxifen. More recently, the feasibility of reporting nonmalignant brain tumors is being investigated. The mechanism also has been used to fund developmental studies of electronic data systems and techniques for improving central registry operations. In addition, SEER Special Studies support development of surveillance methods for larger initiatives that involve multidisciplinary collaborations across several SEER sites, such as the Prostate Cancer Outcomes Study, initiated to address the exponential rise in prostate cancer incidence and associated increase in radical prostatectomy.

lying reasons for these trends has been building. However, major expansion is needed in data collection and resources, including further collaboration and integration with many partners and grantees, to more fully explain why trends occur.

Although the majority of cancer-related health surveillance systems focus on the individual, research has identified the importance of societal, health organizational, and systems factors to effective cancer control. Therefore, surveillance systems that allow adequate evaluation of the effectiveness of our national cancer control program must incorporate measures of all of these elements. The knowledge to be gained includes the following:

- Dimensions (incidence, survival, mortality, health status, quality of life) of the cancer burden and how each dimension changes over

time, in different areas of the country, and in different racial/ethnic groups.

- How factors from risk behaviors through treatment (diet, smoking, screening, treatment, gene-environment interactions) vary over time, geographically, and among racial/ethnic groups.
- How much each factor affects each dimension of the cancer burden, including the lag between changes in risk factors and changes in cancer outcomes.
- Why the cancer burden is changing and what can be done to reduce it.

The remainder of this chapter describes where the NCI currently resides with respect to the points listed above.

### 2.2.1 Dimensions of the Cancer Burden

We currently know how stage-specific incidence and survival change over time and in different racial/ethnic groups in all of the SEER areas. These data are available consistently for all cancer organ sites in all areas of the country served by SEER. Data collection in most of the SEER registries has been occurring in a consistent manner for the past 25 years; therefore, trends in these rates over time are well characterized. CSRP investigators also have developed methods for understanding when an apparent trend is in fact a trend.

The SEER registry areas were selected primarily for their ability to operate and maintain population-based cancer reporting systems and for coverage of population subgroups. Although these areas were not randomly selected under any probability sampling scheme, they were a subset of the U.S. population chosen to be reasonably representative with respect to selected demographic factors for the United States. The SEER registries are regarded as the gold standard for data quality, but they cannot provide information about geographic areas or populations that are not included in the Program.

We know how cancer incidence and survival vary geographically and among racial/ethnic groups in some non-SEER areas of the United States. In 1992, Congress established the National Program of Cancer Registries (NPCR) and authorized the

CDC to provide funds to states and territories to develop or improve their existing cancer registries. In Fiscal Year 1999, with appropriations of approximately \$24 million, the NPCR is supporting 35 states and the District of Columbia to enhance established registries, and 10 states plus 3 territories to develop registries where none existed previously. The CDC will continue to enhance existing state cancer registries and move those states in the planning phase toward data collection activities. As the NPCR-funded cancer registries mature, it is anticipated that their linkage with SEER states and regions will provide full cancer surveillance coverage of the United States. In 1998, data from registries that met specific quality standards were published under the aegis of NAACCR; this report was based on all SEER registries and selected NPCR registries, representing 14 percent and 24 percent, respectively, of the U.S. population (see Figure 4). An overview of cancer registration in the United States is provided in Appendix B.

There is growing recognition of the importance of going beyond the traditional measures of cancer incidence, survival, and mortality to include additional dimensions, such as the impact of cancer morbidity on quality of life as well as measures of patterns of care experienced by cancer patients. Data elements routinely collected by SEER include patient demographics, primary site, morphology, diagnostic confirmation, extent of disease, first course of therapy, and followup for vital status (see Figure 5). The cancer outcomes data routinely collected by SEER registries currently are limited to first course of treatment, date of last followup or death, and underlying cause of death. First course of treatment data are qualitative and characterize broad categories that include site-specific surgery, reason for no cancer-directed surgery, radiation, radiation sequence with surgery, chemotherapy, endocrine therapy, biological response modifier, and other. Data on cancer-related morbidity currently are available for specific organ sites and selected SEER registries collected under Special Studies such as those conducted in collaboration with the National Institute on Aging. In addition to routine collection of this limited information on cancer outcomes, the CSRP is conducting a number of studies that are acquiring data on the cost of cancer treatment and the quality of life experienced by cancer patients, tracking patterns of care in the community

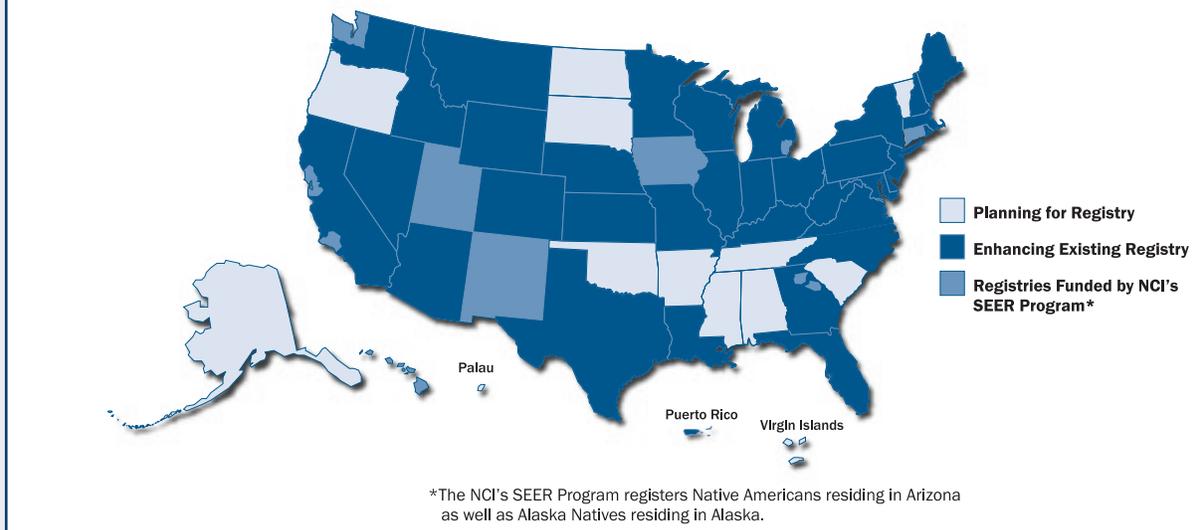
## The National Program of Cancer Registries (NPCR)

In October 1992, Congress established the NPCR by enacting legislation that authorized the Centers for Disease Control and Prevention (CDC) to provide funds to the states and territories to improve existing cancer registries and to plan and implement registries where none currently exist. In Fiscal Year 1998, the CDC supported 45 states, 3 territories, and the District of Columbia for cancer registries, which included enhancement of 36 existing registries and development of 13 new registries. With Fiscal Year 1999 appropriations of approximately \$24 million, the CDC will continue to enhance existing state cancer registries and move planning states toward more thorough data collection

activities. The NPCR registries are expected to collect information on at least 95 percent of the cancer cases diagnosed or treated in their state each year. Timely and accurate data on cancer incidence, stage at diagnosis, first course of treatment, and deaths are collected by the NPCR registries.

When fully operational, the NPCR will collect cancer incidence data on 97 percent (inclusive of SEER) of the U.S. population. The NPCR will improve the representativeness of cancer surveillance data and provide more accurate estimates of cancer incidence for racial/ethnic minority groups.

### National Program of Cancer Registries, 1999



setting of recommended cancer screening and treatment modalities, monitoring the performance of cancer screening at the national level, and documenting patterns of cancer treatment and outcomes in the different health care delivery systems. There is a need to expand this type of resource development and research activity.

In the late 1980s, the CSRP initiated a collaborative effort with the HCFA to create the SEER-Medicare database. This database links the clinical data collected by the SEER registries with claims for health services collected by Medicare for its beneficiaries. These combined datasets can be used for a

broad array of studies including assessing patterns of care for persons with cancer, monitoring use of tests and procedures during the period prior to and following a cancer diagnosis, recording morbidities and comorbidities experienced by cancer patients, and determining costs of cancer treatment. Although this database is a valuable tool for examining these issues in the population over age 65, resources for examining these issues in younger individuals are more limited.

Routine data collection through SEER currently provides information that focuses on events following diagnosis of cancer and on the occurrence of

**Figure 5.**  
Data Characteristics of SEER and NPCR Cancer Registry Programs

Characteristic	SEER Registries	NPCR Registries
Population-based	Yes	Yes
Coverage of U.S. population	14%	Goal is 97%
Cases added annually	160,000	To be determined
<b>Demographic Data</b>		
Age	Yes	Yes
Sex	Yes	Yes
Race	Yes	Yes
Ethnicity	Yes	Optional
Hispanic Origin	Yes	Yes
State	Yes	Yes
County	Yes	Yes
Census Tract	Yes	Yes
Address (GIS)	State/regional registry only	State registry only
Occupation	No	Yes
Tobacco Use	Not required	Yes
<b>Incidence Data</b>	Yes	Yes
<b>Treatment Data</b>	First course only	First course only
<b>Survival Data</b>	Required	Optional
<b>Disease Recurrence</b>	Optional	Optional
<b>Stage-at-Diagnosis</b>	Required	Required
General Summary Stage (GSS)	Can be derived	Required
Extent-of-Disease (EOD)	Required	Optional
Tumor, Node, Metastasis (TNM)	Can be derived	Optional
American Joint Commission on Cancer (AJCC)	Can be derived	Optional
<b>Reporting from Outpatient Sources</b>	Yes	Yes
<b>Internet Availability</b>	Public-use data files	Available through NAACCR

death. Because these data do not provide insight into the quality of life and treatment issues that occur over the entire course of disease, more comprehensive data collection that takes place routinely over the entire longitudinal course of disease for selected cohorts of cancer patients is needed. Currently, the Prostate Cancer Outcomes Study (PCOS) is the only ongoing SEER Special Study that collects primary quality of care data on a longitudinal basis. PCOS was implemented in response to various clinical studies and policy analyses that indicated that the quality of life following treatment for prostate cancer might be a key determinant of the clinical cost-benefit and cost-effectiveness ratios related to screening for prostate cancer. The SEER-Medicare database also supports longitudinal analysis of patterns of care.

### 2.2.2 Monitoring Factors that Affect the Cancer Burden

Except for the demographic risk factors—such as age, race, ethnicity, gender, and first course of treatment—data on factors that affect outcomes are not collected routinely by NCI’s CSRP. However, the

## The SEER-Medicare Database

The SEER-Medicare database is the collaborative effort of the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) cancer registries and the Health Care Financing Administration to create a large population-based source of information for cancer-related epidemiologic and health services research. The creation of the linked files has required matching persons reported to the SEER registries with a master file of Medicare enrollment to determine which persons in the SEER database were eligible for Medicare.

Linking SEER data with Medicare data enhances the potential of two large population-based sources of information. Data from the SEER Program include detailed clinical, demographic, and cause of death information for persons diagnosed with incident cancer who reside in one of the areas participating in SEER. The Medicare data include demographic and eligibility information for all Medicare enrollees

as well as utilization of Medicare-covered health services, which includes all hospital inpatient and outpatient care, physician services, and home health, hospice, and skilled nursing facility use. Therefore, the linked SEER-Medicare data offer an opportunity to examine quality of care issues that occur over the entire longitudinal course of disease (i.e., prior to a cancer diagnosis and during the period of initial diagnosis as well as long-term followup). A number of projects are using the SEER-Medicare data to study the use of health services and patterns of care for specific cancers, to investigate differences in treatment and survival between Medicare Health Maintenance Organization (HMO) and Fee-For-Service enrollees, and to examine the cost of services delivered from diagnosis to death according to phase and stage at diagnosis. SEER-Medicare data have been released to more than 30 extramural investigators.

CSRP utilizes behavioral and other health-related data that are collected routinely by several other government agencies. In addition, the CSRP identifies where new data are needed to answer key surveillance questions and works collaboratively with other government agencies and research institutions in designing and implementing new surveillance activities to collect these data. An important partner in collecting data on risk factors is CDC's NCHS, which is the lead organization within the federal government responsible for collecting, analyzing, and disseminating national health statistics.

The NCHS administers the National Health Interview Survey (NHIS), which is conducted annually in approximately 49,000 households. The NHIS collects health-related data—such as medical conditions, hospitalizations, and visits to the doctor—on all members of the household, which permits an assessment of the factors determining the health practices of the entire household. With leadership from the CSRP, a Cancer Control Supplement to the NHIS was fielded in 1987 and again in 1992. These supplements collected data on cancer risk factors and screening as well as knowledge, attitudes, and beliefs concerning cancer. The CSRP currently is testing the survey instrument for the next Cancer Control Supplement, which will be administered in 2000. The CDC also conducts the Behavioral Risk Factor Surveillance System (BRFSS), which is a continuous, randomly dialed, telephone survey. Because the NHIS and BRFSS are conducted in random samples of the U.S. population, it is possible to know how each factor—from risk behaviors through screening and treatment—varies over time, geographically, and among racial/ethnic groups.

The CSRP works with other federal agencies to develop targeted studies addressing cancer control surveillance issues. For example, in a critical collaboration with the Census Bureau, Tobacco Use Supplements were added to the Current Population Survey to assess progress in national tobacco control. The CSRP also collaborates with CDC's NCHS by developing the Cancer Control Supplement for the National Health Interview Survey. Additional surveys used by the CSRP and other investigators for surveillance research include: the National Health and Nutrition Examination Survey (NHANES) conducted by the NCHS, surveys con-

### The Prostate Cancer Outcomes Study

Prostate cancer is the most commonly diagnosed cancer in the United States, with approximately 200,000 new cases reported each year. There currently is controversy over the appropriate management strategy for clinically localized disease, particularly in elderly men. Recent studies show that long-term survival following conservative management is similar to survival following definitive therapies. To assess the quality of life among men recently diagnosed with prostate cancer, the National Cancer Institute (NCI) is sponsoring the Prostate Cancer Outcomes Study (PCOS). The primary goal of this study is to provide information about health-related quality of life following alternative treatments for clinically localized prostate cancer in a representative patient population treated in community practice. The NCI is working with six SEER registries—Connecticut, Utah, New Mexico, Atlanta, Seattle, and Los Angeles—to collect information from approximately 3,800 men recently diagnosed with prostate cancer. Each registry conducted a mailed survey of a random sample of cases at 6, 12, and 24 months after diagnosis. Participation was voluntary and all information obtained is kept strictly confidential.

Approximately 3,500 men participated in this repeated survey, which included questions on urinary, bowel, and sexual function as well as general health status and concerns about prostate cancer or its treatment. Because quality of life is expected to differ depending on tumor characteristics and medical treatment, this information was obtained from the patients' medical records. The results of the PCOS will be published in peer-reviewed medical journals so that physicians and researchers around the country will benefit from the improved understanding of the impact of prostate cancer on men's everyday lives.

ducted by the U.S. Department of Agriculture (USDA)—the Continuing Survey of Food Intakes by Individuals (CSFII) and the Nationwide Food Consumption Survey (NFCS), and the Youth Risk Behavior Survey administered by the CDC.

A number of data resources have been developed by the CSRP to facilitate cancer surveillance research. To evaluate the effect of NCI's 5 A Day Program for

## The Cancer Control Supplement of the National Health Interview Survey

The National Health Interview Survey (NHIS) is a continuing annual nationwide survey of households of the civilian, noninstitutionalized population conducted by the National Center for Health Statistics and the U.S. Bureau of the Census. The survey is administered in person in about 49,000 households each year. In 1987 and 1992, Cancer Control Supplements were administered as part of the NHIS to an adult sample of persons age 18 years and older to determine knowledge, attitudes, and use concerning cancer screening modalities and risk factors. For both supplements, black and Hispanic households were oversampled. In 1987, the Cancer Control Supplement was administered to 22,043 respondents and to 12,035 in 1992. The Supplements have included questions on tobacco usage to obtain data on smoking, passive smoking, smoking in the workplace, attitudes regarding smoking, and use of smokeless tobacco. In addition, the Supplements have included a 60-item food frequency questionnaire designed to capture intake over the past year of nutrients related to cancer prevention and control. Also included in the Cancer Control Supplements were questions regarding alcohol consumption, hormone use (such as birth control pills or estrogen replacement therapy), occupa-

tional exposures, cancer survivorship (issues related to quality of life), and knowledge and attitudes pertinent to diet and health. The Supplements also captured data on cancer screening, including frequency of mammogram, breast self-exam, Pap smear, digital rectal exam, blood stool test, proctoscopic exam, and chest x-ray; location of examination; accessibility of medical facility or physician for screening; type of medical insurance; and attitudes and knowledge regarding screening.

The next Cancer Control Supplement, which will be fielded in 2000, currently is being developed and tested. The questions will include those described above (except for the chest x-ray question and the section on occupational exposures) as well as a set of questions about prostate specific antigen (PSA), updated questions on colorectal cancer screening (FOBT home test and endoscopy), and sections on family history of cancer and genetic testing for cancer risk. The dietary section was modified to include brief screener questions designed to capture information on intake of fruit and vegetables, percentage of energy from fat, and intake of fiber. This modification will allow the next Cancer Control Supplement to be administered to the entire NHIS sample.

Better Health, the CSRP worked collaboratively with the program staff on baseline and followup surveys conducted in 1991 and 1997, respectively. Anticipating the need for national data on the quality of community-based mammography, the CSRP conducted the 1992 National Survey of Mammography Facilities, which provided the first nationally representative profile of screening mammography in the United States, including pricing, organizational characteristics, systems, and clinical performance when biological markers of risk are identified and validated. Other resources developed by the CSRP include the NCI Black White Study, the SEER Patterns of Care studies, the SEER-Medicare database, the PCOS, and the linkage of SEER data to census tract level data on the demographic, socioeconomic, and health resources characteristics of small areas. Appendix C provides a brief description of the surveys and data sources used by the CSRP.

### 2.2.3 The Relationship of Risk Factors, Screening, and Treatment to the Cancer Burden

Our knowledge about how risk factors, screening, and treatment affect the cancer burden is limited. How and whether patterns and trends in risk factors may be responsible for observed trends in cancer outcomes cannot be determined readily using existing data sources or statistical methods. Information about the effects of changes in behaviors and other factors on dimensions of the cancer burden comes primarily from randomized trials and case-control studies. The data collected and managed by the CSRP provide an infrastructure to support such studies. However, data from these more controlled settings do not provide an accurate estimate of how these factors may be influencing the population

cancer burden. Knowledge about the relationships among risk factors, cancer incidence, and outcomes is imperfect and there is limited understanding of the lag between changes in risk factor behavior and their effects on cancer outcomes.

Data within existing health surveillance and care delivery systems often are not sufficient to answer specific questions regarding whether cancer control behaviors and practices found to be beneficial in controlled clinical trials actually result in the same degree of benefit once translated into community practice. Even when *efficacy* estimates are available, *effectiveness* in a defined population may be unclear. For example, controlled clinical trials of breast cancer screening by mammography indicate a 25 to 30 percent reduction in mortality; these results, however, may not be generalizable to the U.S. population for several reasons. These include possible differences in the population groups receiving screening, lower accuracy of screening mammography in the community, and lower compliance with diagnostic followup and treatment in community practice.

To examine the effectiveness of mammography screening in community practice and to determine if the benefit in the reduction in mortality estimated from clinical trials is occurring at the population level, the CSRFP developed the Breast Cancer Surveillance Consortium (BCSC), a consortium of eight sites supported by cooperative agreements with the NCI. In addition to collecting core data on screening behaviors, the BCSC has supported special research projects addressing issues such as cost effectiveness, quality of life, and biological differences in screen-detected versus nonscreen-detected cancers. Although the BCSC has focused primarily on the effectiveness of screening mammography, it also offers the opportunity to examine the diffusion of mammography in defined populations and has included the collection of limited risk factor data. With the recompetition of the BCSC, new data linking direct assessment of mammography results to cancer outcomes will become available. Similar questions arise for several other major cancers, but cannot be addressed with existing data resources.

Currently, health surveillance systems that allow linkage of variation in risk factor and screening measures to cancer outcomes in defined populations are limited. This fact has led to several new CSRFP extramural research initiatives. For example, the

### The Breast Cancer Surveillance Consortium (BCSC)

The National Cancer Institute (NCI) initiated a number of pilot studies in 1990 to determine if screening in community practice resulted in breast cancer mortality reductions comparable to those demonstrated in clinical trials. With a mandate from the 1992 Mammography Quality Standards Act, the NCI established the Breast Cancer Surveillance Consortium (BCSC) in 1994 to evaluate population-based screening mammography in the United States. The three major objectives of the BCSC are to: (1) enhance our understanding of breast cancer screening practices in the United States through an assessment of the accuracy, cost, and quality of screening programs and the relation of these practices to changes in breast cancer mortality or other shorter-term outcomes, such as stage at diagnosis or survival; (2) foster collaborative research among Consortium participants to examine issues such as regional and health care system differences in providing screening services and subsequent diagnostic evaluation; and (3) provide a foundation for clinical and basic science research that can improve understanding of breast cancer etiology and prognosis.

The geographic, rural, and minority representation within the BCSC was expanded in 1995. Currently, eight sites throughout the United States are participating in the Consortium. The NCI issued a Request for Applications (RFA) in early 1999 that will provide funding for 9 to 11 data collection sites as well as a Statistical Coordinating Center. This RFA will continue to support current research efforts, including the ability to track the diffusion of new screening technologies, obtain more detailed risk factor data in special research projects, and expand the diversity of populations represented within the BCSC. By 2000, it is estimated that the BCSC database will contain information on nearly 3.2 million screening mammographic examinations and more than 24,000 breast cancer cases.

Health Maintenance Organization Cancer Research Network (HMO CRN)—a consortium of 10 large, nonprofit, research-oriented HMOs—is funded as a cooperative agreement with the NCI to serve as a laboratory for population-based cancer control research.

## The HMO Cancer Research Network

The purpose of the Cancer Research Network (CRN) is to encourage the expansion of collaborative cancer research among health care provider organizations that are oriented to community care, are able to take advantage of existing integrated databases that provide patient-level information relevant to cancer control and cancer-related population studies, and have access to large, stable, and diverse populations. In 1999, the first CRN project was awarded to the HMO Cancer Research Network, which is a consortium of researchers affiliated with 10 major nonprofit health maintenance organizations (HMOs). The aim of this Network is to establish a population laboratory for high-quality research to determine the effectiveness of cancer control interventions that span the natural history of major cancers among diverse populations and health systems. The objective is to identify those patient, treatment, and delivery system factors associated with differences in outcomes or costs. In collaboration with the

National Cancer Institute (NCI), the HMO Cancer Research Network will:

- Develop an administrative infrastructure to support research collaboration, data quality and integrity, and development of methods and approaches to increase the participation of managed care patients in NCI-approved clinical trials.
- Study the efficacy, reach, and quality of delivery as well as adherence to smoking cessation programs as delivered in HMO practice settings.
- Examine late-stage breast and invasive cervical cancer cases to elucidate those patient, provider, and system factors that contribute to preventing advanced disease.
- Study the effectiveness of the commonly used strategies of frequent mammography or prophylactic mastectomy, to prevent fatal breast cancer among women who are at increased risk.

### 2.2.4 Why the Cancer Burden is Changing and What Can Be Done to Reduce It

The CSRP has initiated a number of “Surround SEER” activities to evaluate patterns and trends in cancer-related risk factors, health behaviors, and health services and to determine the influence of these factors on trends in cancer incidence, morbidity, mortality, and survival. These surround SEER studies are conducted by the CSRP as well as extramural researchers. CSRP’s Applied Research Branch currently carries out research and evaluation activities in three existing sections—Health Services and Economics, Surveillance Modeling and Methods, and Risk Factor Monitoring and Methods—and a new section addressing outcomes research within clinical trials and at the population level has been established. CSRP investigators are developing methods to estimate cancer prevalence and lifetime and age-conditional risks of developing and dying of cancer. In addition, the CSRP is developing methods and models that account for lags in effects of changes in risk factors and screening on incidence and outcomes. Information on these and other efforts of the ARB is available on the Internet at <http://www-dccps.ims.nci.nih.gov/ARB/>.

Currently, it is possible to examine patterns and trends in nationally representative samples, and in some selected national and regional samples, by a number of demographic measures. These include gender, age, and major racial/ethnic groups (i.e., white, African American, Asian American, Hispanic) as well as various socioeconomic measures such as education and very broad measures of income. Interrelationships among behaviors and practices can be examined in some health surveillance systems to a limited degree. Trends in cancer incidence, stage-at-diagnosis, survival, and mortality can result from a number of factors. For example, a marked trend in cancer incidence could be a result of either changes in exposure to risk factors or changes in the use of cancer screening services. SEER data have been used by a number of investigators to explain the relationships between behaviors and screening practices and cancer outcomes. One example is the investigation of the rise in breast cancer incidence from 1982 to 1987. The relationship between the increase in screening by mammography and the rise in breast cancer incidence was examined to distinguish such an effect from an epidemic of breast cancer. The investigation indicated that patterns of breast cancer incidence are generally close to that predicted by the increase in screening.

# 3 The Strategy: Where Do We Go From Here?

The SIG members strongly endorsed the current data collection and research methods development activities of the CSRP. They concurred with the CCPRG that the surveillance program is critical to NCI's ability to control cancer, and agreed that investment should be targeted to activities that will help explain patterns and trends in the cancer burden. It is essential that support for the SEER Program be continued to ensure that the NCI can accurately track changes in cancer rates and provide data to support epidemiologic research. The SIG also noted the importance of expanding the surveillance research program to allow the CSRP to fully realize its potential as a resource for genetics and outcomes research.

The highest priority for investment in CSRP activities is in the area of explaining observed rates and trends in dimensions of the cancer burden. The CSRP currently reports on cancer trends and has begun to use many different data sources and models to explain why trends are changing. Additional data resources and methods are needed to provide more comprehensive answers to these questions. Members of the SIG agreed that the CSRP should work with interdisciplinary scientists throughout the NCI and in the extramural community to provide answers to questions about the cancer burden. As detailed in the previous chapter, decision makers need to know the dimensions of the cancer burden and how they are changing over time, in different areas of the country, and in different racial/ethnic groups; how risk factors (e.g., diet, smoking, screening, treatment, gene-environment interactions) vary over time, geographically, and among racial/ethnic groups; how these factors influence each dimension of the cancer burden; and what opportunities exist to reduce the burden.

The SIG members considered the data and methods that would be needed to answer such questions and concluded that the CSRP is taking important steps to address these needs. Data collection is needed both for cancer patients and the general population (the general population includes individuals at risk for developing cancer). Because the questions and issues that need to be assessed are different for these two distinct groups, data collection efforts and the best approaches to accomplishing these also are different. Risk factor and screening data must be collected from the general population prior to the development of cancer; data on treatment, however, must be obtained on cancer patients. Data on quality of preventive care, including behavioral interventions and screening, are needed. In the future, as screening becomes more widespread and a larger proportion of cancer is detected in early stages, data on prognostic factors for cancer recurrence, including tumor markers and data on preventive health behaviors, will become increasingly important to collect in cancer patients.

The SIG believes that a number of opportunities exist to enhance NCI's surveillance program and that modest investments in appropriate areas could have a profound impact on the ability of the NCI and the extramural research community to explain observed cancer rates and trends. Through a consensus-based process, the SIG members identified 12 cancer surveillance research opportunities, which were organized within five overarching implementation priorities. These research opportunities are the focus of this chapter. For each of the opportunities, the SIG has identified an estimate of the approximate level of investment required and a time frame for initiating the research. "High" costs are anticipated for investments of more than \$3 million an-

nually, “low” costs for investments of less than \$500,000 annually, and “moderate” costs for the intermediate range. Determination of exact costs

will have to await the development of specific concepts, special study initiatives, and joint discussion with partners for collaborative initiatives.

## Priority Area 1: Expand the scope of surveillance research through additional data collection and methods development.

Within this priority area, the SIG identified the following four critical needs:

- Data on additional measures (e.g., patterns of care, quality of life) of cancer burden beyond incidence, survival, and mortality should be collected within established population-based cancer registries to fully assess the Nation’s cancer burden.
- Representative data about risk factors and screening should be collected from defined populations where SEER-quality cancer registration occurs to improve our understanding of which cancer risk factors and screening behaviors lead to improved cancer outcomes and to track trends in these factors and behaviors over time.
- Methodologic research is needed to develop appropriate tools and methods for managing and analyzing these additional data on measures of the cancer burden and related risk factors.
- The utility of geographic information systems (GIS) as an innovative addition to the cancer surveillance infrastructure should be investigated.

### Research Opportunity 1

**Support the collection of data on patterns of care, health status, morbidity, and quality of life as well as cohort studies of newly diagnosed registered cancer patients for the purpose of documenting levels and trends in these parameters.** *(The cost for this effort is expected to be high; work should be initiated within the next 1-2 years.)*

The cancer burden in the United States currently is characterized by incidence, survival, and mortality. There is growing recognition of the importance of additional measures that go beyond these “core” measures. The CSRIP has explored these issues and

developed a number of other measures, including those derived from the basic data such as cancer prevalence and lifetime risk of dying from cancer. However, a state-of-the-art cancer surveillance system should report on cancer morbidity as well as mortality. Thus, data on patterns of care and cancer outcomes beyond incidence, survival, and mortality are needed. Research is needed to determine the best measures of patterns of care, health status, morbidity, and quality of life. Likewise, support is needed for the collection of these measures in incidence cohorts within established population-based cancer registries. A specific goal of this research would be to better measure the cancer burden (including quality of life, health status, and morbidity) and to measure the medical interventions that affect survival within stage at diagnosis, morbidity, and mortality. Research should focus initially on the quality of treatment (including entry into controlled clinical trials) and quality of life among cancer patients from the time of initial diagnosis through long-term survival and end-stage care.

There are some data elements that currently are not collected, either routinely or through special studies, that would greatly enhance our ability to monitor important aspects of quality of life and quality of care. Routine data collection through SEER currently provides information that focuses on events following diagnosis of cancer and on the occurrence of death. Although this information is crucial, it does not provide insight into the quality of life and quality of care issues that occur over the entire course of disease and treatment. Therefore, more comprehensive and routine data collection, which takes place over the entire longitudinal course of disease for selected cohorts of cancer patients, will add significantly to our understanding of the overall burden of cancer and how various cancer control interventions might reduce this burden.

The SIG recommends that this research be implemented by developing a series of site-specific cancer cohorts for major cancer sites (i.e., breast, colon, prostate, and lung). A model for such research is the Prostate Cancer Outcomes Study, which was supported through a SEER Special Studies mechanism. PCOS was implemented in response to various clinical studies and policy analyses that indicated that the quality of life following treatment for prostate cancer might be a key determinant of the clinical cost-benefit and cost-effectiveness ratios related to screening for prostate cancer. In this study, a cohort of approximately 3,500 men with newly diagnosed prostate cancer has been followed for several years with repeat surveys addressing health-related quality of life and care patterns.

The SIG suggests that SEER Special Studies or Requests for Applications (RFAs) could be appropriate mechanisms for extending the PCOS model to other major cancer sites (i.e., breast, colon, and lung). Applicants would collaborate closely with CSRP staff to identify and initiate cohorts studies of newly diagnosed patients. Existing SEER-quality cancer registry sites would be eligible and encouraged to apply. All grantees would agree on a minimum set of core variables to be collected consistently for all individuals being followed, according to quality standards established under the leadership of the CSRP. Individual sites could collect additional information to meet the needs of special studies. Participating research sites should have an efficient system in place for rapid case ascertainment. The presence of a computer-assisted telephone interview (CATI) system that could be used for rapid case ascertainment of a random sample of newly diagnosed patients as well as for conducting interviews on a routine basis would facilitate these studies. The CSRP should consider using Program Announcements (PAs) to support similar research for less common cancer sites.

### Research Opportunity 2

**Support the collection of risk factor and screening data in defined populations, particularly those covered by high-quality cancer registration.** (*The cost of this effort is expected to be high; work should be initiated with the next 1-2 years.*)

One objective of the CSRP is to explain trends in dimensions of the cancer burden in terms of

lifestyle and behaviors such as smoking, diet, physical activity, and participation in screening. To accomplish this, data are needed on patterns and trends in risk factors and screening over time in populations defined by geography, race/ethnicity, income, and other important demographic, social, and biobehavioral characteristics influencing cancer control. The key is to obtain representative data about risk factors and screening from defined populations where SEER-quality cancer registration occurs. Data on patterns and trends in these factors can be drawn from multiple existing health surveillance systems, some of which will need to be expanded to address current questions. Data on cancer outcomes then can be linked to data on cancer mediators. To make these linkages with other data sources possible, the completeness and accuracy of specific data items currently collected by SEER must be improved, including cause of death and address at time of diagnosis.

To explain trends in cancer outcomes in terms of behaviors, measurement of trends in these behaviors is required. Data on individuals at risk for cancer before they become cancer patients must be collected to understand how *effectiveness* in a defined population differs from *efficacy* in trials and case-control studies. Direct collection of data at the community level is necessary to address this issue and advance understanding of the practice and quality of cancer control at the population level.

The SIG recommends this research be accomplished by: (1) initiating in-depth studies to link risk factor and screening data directly to cancer outcome data, and (2) improving the capacity to monitor patterns and trends in cancer-related risk factors. The NCI could support expansion or development of selected local and national health surveillance systems and databases with geographically specific data. The CSRP also could support RFAs or PAs for targeted innovative research efforts to collect risk factor and screening data on defined local populations linked to SEER-quality cancer registry systems. Cooperative agreement mechanisms have been used by the CSRP to facilitate coordination of research and the development of comparable data across diverse research groups, and such mechanisms should be continued to support this research.

There is a need for research on approaches to link risk factors and screening behaviors to outcomes

through the expansion of existing national health surveillance systems (e.g., the Current Population Survey, the National Health Interview Survey, and the Behavioral Risk Factor Surveillance System). Partnerships with these surveillance systems will be valuable with respect to the collection of risk factor, screening, and behavioral data. To include new tracking systems to monitor organizational and psychosocial factors that facilitate or impede behavioral lifestyle changes in youths and adults, partnerships should be developed with institutions inside and outside the government. Existing partnerships with the CDC (including the NCHS) and the HCFA should be strengthened, and new partnerships with other organizations should be initiated. Methods to link the databases of these institutions to the data available from SEER are needed. Standardization of questions for data collection and database development are key to facilitating these linkages, which will produce information about the relationships between risk factors and outcomes. It will be necessary to work with relevant health agencies and partners regarding the adaptation and development of local and national health surveillance systems for cancer surveillance, leading perhaps to interagency agreements or contracts for data collection.

Extramural investigators would collaborate closely with CSRP staff to identify or initiate risk factor surveys in the defined population covered by a cancer registration system. Specific goals of this collaborative effort would include, but not be limited to, measuring risk factors such as family history, environmental exposures, behaviors (e.g., smoking, diet, physical activity), screening, treatment, and cancer biology. Applicants might propose to link risk factor data with cancer data, link screening data with cancer data, and/or develop methods for measuring the factors that affect cancer burden. Existing sites with SEER-quality cancer registration would be eligible and encouraged to apply. Applicants would agree to collect a set of core variables consistently, according to quality standards established under the leadership of CSRP staff. These efforts might involve de novo surveillance in an area or alternatively, an applicant might propose to partner with an agency that has an ongoing risk factor survey in its state, and work with that organization to meet the established standards with respect to comparable definitions of the content and/or analytical methods.

Research on the relationship between screening practices to cancer outcomes should be implemented by developing consortia modeled after the BCSC, in which mammography screening data are linked to cancer surveillance data. Specifically, the development of a Colorectal Cancer Surveillance Consortium is a high priority. Grantees would agree on a set of core questions that would be asked of all individuals; however, participating sites might collect additional self-report, medical record, or biologic data to meet the needs of special studies. The consortium would collect limited patient-level risk factor data and detailed screening and cancer outcome data as well as information on health care providers. An important element of such efforts includes collecting data on organizational and system factors that might influence screening performance. For example, the BCSC sites are collecting and linking data on screening mammography from radiologic practices, information collected in the medical record at the time of screening mammography (limited data on past screening history, risk factors, and symptoms at the time of mammographic examination), followup information regarding repeat radiologic examinations and subsequent diagnostic surgical procedures, and pathologic data and cancer outcomes from either pathology laboratories or cancer registries.

### Research Opportunity 3

**Develop research methods to measure dimensions of the cancer burden and factors affecting the burden as well as methods to explain patterns and trends in cancer rates.** (*The cost for this effort is expected to be moderate; work should be initiated within the next 1-2 years.*)

As data on additional measures of the cancer burden and related risk factors are collected to track the Nation's progress against cancer, appropriate tools for managing and analyzing these data will be needed. Methodologic research will be necessary to develop these tools and methods for better characterizing cancer-related risk factors and the cancer burden, explaining why that burden is changing, and determining what can be done to reduce it.

The ability to interpret cancer trends is dependent on high-quality data and innovative statistical methods to allow inference based on diverse sources of data. The quality of methods for ascertainment of

exposures and other measures, such as quality of life and family history, dictate the quality of data collected.

Improved methods to assess dietary intakes to determine the relationship between intake and disease risk are needed for dietary surveillance to determine how the population's intake compares to recommendations, how intake relates to other factors, and whether diet is improving over time, and for epidemiologic research to determine the relationship between intake and disease risk. CSRP's public-use dietary assessment instruments, software programs (e.g., DietSys) for nutrient analysis of these instruments, and information on validated and calibrated dietary assessment tools (e.g., the Dietary Assessment Manual) are valuable resources that should be used by the extramural research community to support this research.

The methodologic issues surrounding dietary assessment include every phase of research, from instrument development and data gathering to analysis. Among the most pressing issues are estimating the distribution of usual intakes in the population, and its corollary, estimating the usual intake for each individual in the sample. Usual intake refers to the long-run average intake of a dietary component and represents the theoretical variable of interest for most dietary research. Other critical needs are the development of appropriate assessment tools for specific populations, the development of electronic versions of assessment instruments, and validation research of self-report diet with biologic measures of diet. An example is the use of serum and urinary measures of nutrients, and doubly-labeled water for evaluation of energy. The CSRP should continue to improve the usability of the major federal food consumption surveys, the CSFII conducted by the USDA and the NHANES conducted by the NCHS, and address the methodologic issues resulting from the Year 2000 merger of these two surveys.

Methods to characterize lifetime weight patterns and measures of current and past physical activity are needed because energy balance over critical stages of the life cycle may be an important risk factor in the development of some cancers. It will be important to determine what measures of physical activity are most relevant to track for selected cancer sites, to develop better methods for their assessment,

and to determine how physical activity is associated with a number of cancer-related health behaviors including weight, smoking, and alcohol use.

Measures of tobacco use in at-risk populations (e.g., youth) are needed as well as methods to assess environmental, economic, and legislative policies that influence tobacco use; biologic measures of nicotine exposure; and evaluation of the impact of the ASSIST program on tobacco use. These efforts will provide more complete data on tobacco use and tobacco control policies to assist policymakers and health care planners in developing more effective strategies for tobacco control. Future research priorities include methodologic work to improve biologic and self-report measures of tobacco use, particularly in at-risk populations such as youth, and development of measures of home smoking policies and passive smoke exposure.

The best measures of quality of care that can be applied systematically for the Nation and reflect change over time should be identified. In addition to patient level measures of quality of care, health system and organizational measures of quality of care are needed. Such measures might be incorporated both in clinical trials and in cancer cohorts described earlier.

Finally, there is a critical need to develop methods in statistical surveillance analysis. The CSRP has built a strong program to develop methods to solve quantitative problems in cancer surveillance and to study the impact of cancer control interventions on the cancer burden. The development of many of the methods and programs supported by the CSRP has occurred through a variety of mechanisms. With the increasing use of these quantitative methods, collaborative exchange among investigators would advance ability to compare findings from these diverse approaches. Therefore, the CSRP is developing a proposal, the Cancer Intervention and Surveillance Modeling NETwork (CISNET), which is an effort to develop a consortium of centers to: (1) support modeling research of population-based trends in risk factors, screening, and treatment related to cancer outcomes; (2) help design, interpret, and extrapolate screening and prevention studies; and (3) evaluate cost and health effects of specific interventions.

The SIG recommends research to improve methods for ascertaining a number of cancer-related mea-

tures, including, but not be limited to, developing methods for measuring quality of life, health status, morbidity, family history, environmental exposures, behaviors (e.g., smoking, diet, physical activity), screening, treatment, and cancer biology as well as methods and models for relating variables and predicting outcomes.

The SIG supports approval of an RFA for CISNET as an appropriate and timely approach for stimulating extramural research in this area. The extramural investigators should collaborate closely with CSRP staff to develop research methods and models for measuring dimensions of the cancer burden and factors affecting the burden as well as methods and models for explaining trends and patterns in cancer incidence and mortality.

#### Research Opportunity 4

**Explore the feasibility and utility of employing geographic information systems for geocoding surveillance data and reporting geographic relationships among screening measures, risk factors (including environmental exposures), and improved cancer outcomes. Methods need to be developed for assuring data confidentiality.** *(The cost for this effort is expected to be moderate; work should be initiated within the next 1-2 years.)*

Research is needed on the utility of geographic information systems (GIS) as an innovative addition to the cancer surveillance infrastructure. It is anticipated that GIS will be an important tool in creating and facilitating the use of a number of data resources for cancer surveillance. A wealth of geocoded data currently is being collected by a large number of groups for a diverse set of purposes. These data could be useful for examining cancer etiology and control, but they currently are not linked to individual-level data on cancer. Epidemiologic examination of the etiology of cancer has been limited to characteristics that could be reported accurately by the individual or measured by investigators. Systems have been developed that allow precise spatial description of phenomena that are not reported accurately by individuals, and rapidly expanding technology allows accurate location of individuals using the same spatial descriptors. These recent advancements will facilitate the linkage of potentially important risk factors to cancer outcomes in defined populations.

SEER data can be used effectively to track patterns and trends in cancer incidence, stage at diagnosis, survival, mortality, and some patterns of care by locality, race, gender, and age. To explain these patterns and trends, researchers need detailed information on socioeconomic status, neighborhood characteristics, and environmental factors. This information would be much more helpful if it could be linked closely to the residences of SEER cases. By capturing county, census tract, and the coordinates (latitude and longitude) of where cancer patients live in the SEER system, the number of linkages with other geographically based data systems becomes maximized. To support cancer surveillance efforts, there is a need for geocoded data on socioeconomic status, residential neighborhood characteristics (including health, business, and social services), and environmental factors. In addition, validation of geographic coding is needed to facilitate quantification of the relationships among risk factors, screening processes, and cancer outcomes for defined geographic areas.

Health data that are independently geocoded to an x,y-position on the earth are more useful than data that have been placed in arbitrary or administrative spatial units, such as health districts or blocks. This is especially the case when other information to which a researcher expects health data may be related is geocoded to other spatial units, making necessary the co-registration of health and other information. The U.S. Bureau of the Census "TIGER" files are being used increasingly for automatically assigning geographic codes to health records. GIS brings efficiencies and qualitative improvements to spatial analysis. Furthermore, GIS methods potentially allow linkages for the entire country for incident and fatal cancers and thus can be used to help explain local phenomena and trends. GIS could evolve into an important tool in cancer prevention and control. However, it is critical that mechanisms for protecting confidentiality be developed to maximize the utility of this technology. Spatial aggregation, which has been the standard method for preserving confidentiality of geographic data, will not suffice for health-related GIS activities. The SIG recommends that research be conducted to develop alternative methods to guard the privacy of health records incorporated in GIS-based geographic analysis.

The CSRP should consider developing and testing a health-related GIS. In-house collaborations with SEER Principal Investigators and data managers could be undertaken to explore the potential for improving the accuracy and extending the geocoding of SEER data as well as matching SEER with census data.

The SIG recommends that both in-house and extramural research into health-related GIS spatial statistics and methods development be encouraged and supported. For optimal GIS use, there must be GIS-compatible software for linking data and integrating spatial/statistical methods with GIS. To create the software products needed for further development in cancer control, potential users would benefit by participation in the process of GIS software development, design, and adaptation for cancer research. Statisticians will need to interact and collaborate as part of the multidisciplinary GIS team.

Epidemiologists usually are not trained to think spatially and may not realize or comprehend the potential usefulness of GIS. Interaction of epidemi-

ologists with GIS-based scientists of diverse disciplines will promote interdisciplinary communications and collaborations for enhancing epidemiologic approaches. To this end, the SIG recommends that GIS-focused workshops and conferences be sponsored by the CSRP. These workshops would involve multidisciplinary groups for exchange of knowledge bases and rethinking of approaches to research problems in cancer epidemiology (i.e., reinvention of GIS). Establishing liaisons with other agencies, such as the U.S. Geological Survey's (USGS) Health Committee scientists who are working to identify natural resource data and information relevant to public health, should be considered. GIS/cancer epidemiology sessions might be added to future professional and organizational meetings (e.g., International Society of Environmental Epidemiologists, Agency for Toxic Substances and Disease Registry (ATSDR) conferences on "GIS in Public Health"). Communications also could be increased by use of Intranet/Internet and the bimonthly electronic newsletter, *GIS News and Information*, which is published by the CDC and the ATSDR.

## Priority Area 2:

### Expand the scope of surveillance to improve the representativeness of cancer burden estimates.

The SIG identified the following three needs within this priority area to improve national cancer burden estimates:

- Partnerships among the various agencies involved in national cancer surveillance are needed to expand SEER-quality cancer registration within racial/ethnic minority and medically underserved populations.
- National estimates of cancer risk, incidence, morbidity, and mortality should be developed that are representative of the entire Nation.
- A strong alliance must be forged among the agencies responsible for national surveillance efforts to develop and implement a plan to achieve complete and comprehensive cancer registration for the country.

## Research Opportunity 5

**Expand NCI's surveillance program to improve representation of ethnic minority and underserved populations.** (*The cost of this effort is expected to be high; work should be initiated within the next 1-2 years*)

The most critical component of a cancer surveillance system is a comprehensive description of the cancer burden itself, including incidence, survival, and mortality from every form of cancer. National estimates of these cancer rates should be representative of the Nation as a whole, and local area estimates should be both available and comparable to each other and to national estimates. Although efforts have been made in the past to add under-represented segments of the population—particularly

African Americans, Hispanics (Mexican Americans), Asians, and Pacific Islanders—to the SEER Program, SEER is not a fully representative sample of rural African Americans, Hispanics from Caribbean countries, American Indians, and residents of Appalachia and other rural areas, especially those of lower socioeconomic classes.

To ensure that racial and ethnic minorities as well as medically underserved populations are adequately represented both nationally and locally, partnerships among the various agencies involved in national cancer surveillance are needed to expand SEER-quality cancer registration within these populations in the United States. Such expansion will provide a better understanding of the cancer burden in different areas of the country and in different racial/ethnic and medically underserved groups.

The recommended expansion should be accomplished by working with CDC's NPCR and NCHS, the NAACCR, and states to collect and report SEER-quality incidence, mortality, and survival data on one or more of these populations. It is envisioned that the NCI, the CDC, local administrative units, professional organizations, other federal agencies, and private organizations involved in standard setting and approval programs will form a mutually beneficial partnership to develop this resource for the country. The CDC and the NCI should develop mechanisms to provide funding to local areas to collect data in a standard format that can be used to generate local and national estimates of the cancer burden, in special populations as well as for the population as a whole. It probably will take several years to move from the current system to the comprehensive, representative system that is envisioned by the SIG. However, monitoring of cancer incidence and mortality in local areas offers the benefit of early identification of trends in incidence and mortality that are not usually distributed uniformly throughout the country.

### Research Opportunity 6

**Explore methods for developing improved national estimates of the cancer burden.** (*The cost for this effort is expected to be low; work should be initiated this year.*)

National estimates of cancer risk, incidence, morbidity, and mortality should be representative of the entire Nation. The 14 percent of the U.S. population covered by SEER registries does not represent completely the national cancer burden in a statistical sense. Short of an eventual goal of complete cancer registration in the United States, research is needed to develop and consolidate the following two complementary approaches for improving national cancer burden estimates: (1) sampling methods and estimation techniques that, with the appropriate data, would generate a probabilistically based estimate of the cancer burden; and (2) techniques for modeling national rates based on SEER and other population-based registries.

Probability sampling should be considered for selecting a national sample of cancer incidence cases and making comparisons among the viable sample designs with respect to cost and magnitude of sampling and nonsampling errors. For example, the feasibility of a sample design that expands the existing SEER Program by augmenting it with a random sample of SEER-quality cancer registries and with a random sample of areas of the country that do not have well established registries could be explored. These existing registries and areas without well established registries could be stratified according to cost and data quality considerations where an optimal stratified sampling scheme would be used to select the augmented sample of registries and areas. All cancer incidence cases could be utilized from the augmented sample, as is done in the SEER registries. In the implementation of this design, there would be a number of decisions to be made, such as: (1) how to form into distinct areas those parts of the United States that do not have well established cancer registries; (2) what kinds of sampling should be used to sample these areas within strata, e.g., simple random sampling or sampling proportional to population size; and (3) how to sample the existing non-SEER registries.

The use of model-based estimates is a relatively inexpensive method for improving national estimates of the cancer burden. One method for obtaining model-based estimates of the cancer burden involves regressing demographic factors from existing registries on cancer rates and applying these regression

equations to the demographic profile of the United States. Another method is to “back-calculate” incidence from cancer mortality.

The CSRP should consider the feasibility of these two different approaches as methodologies that may be used in combination to produce data on various aspects of cancer incidence and survival. Pilot studies could be initiated to investigate potential sample designs for selecting national probability samples of cancer incidence cases.

### Research Opportunity 7

**Work with partners to develop a National Cancer Surveillance Plan.** (*The cost of this effort is expected to be low; work should be initiated this year.*)

There currently are several national efforts in cancer surveillance, each of which was designed for a specific purpose (see Figure 6). By combining data from the SEER registries with that from the NPCR registries, cancer incidence data soon may be available for the entire United States. However, the reporting periods for the non-SEER registries currently lag behind those of the SEER registries and the data quality (completeness and accuracy) is variable. The eventual goal of complete cancer registration in the United States can be achieved only by an effective working partnership among the agencies responsible for these surveillance efforts and their primary end users. The SIG recommends that the CSRP work with these partners in the context of the NCCCS to forge a successful alliance to develop and implement a plan to achieve complete and comprehensive cancer surveillance for the country.

Under the auspices of the NCCCS, agencies and individuals involved in cancer surveillance would be invited to present their goals for a national surveillance system. Participants would be asked to share their visions and approaches for obtaining complete cancer registration in a defined population, collecting data on lifestyle and screening behaviors in the same population, and following newly diagnosed

cancer patients for outcomes. The goal would be to sustain a dialogue with partners and with extramural investigators that would nurture a shared vision for an exemplary cancer surveillance system for the nation and development of a national cancer surveillance plan.

The national plan should include a strategy for:

- Identifying and evaluating a process to facilitate registration of patients diagnosed and managed in ambulatory care settings, including the physician's office and HMOs.
- Coordinating electronic resources for cancer registry data, including the development and dissemination of information via the Internet to support data collection and communications.
- Establishing a standardized pathway and ultimately computer algorithms, to align common coding systems.
- Exploring the potential for uses of national and state-level data by hospital clinicians and the general public.
- Establishing a joint committee focusing on quality control issues for cancer registries that will ensure data quality and comparability among cancer registries.
- Identifying and conducting collaborative patterns of care studies.
- Evaluating the impact of new standards promulgated by oversight organizations.

To achieve the goal of a national surveillance system, there must be continued coordination among the organizations involved in cancer surveillance, a more visible and specific commitment of NCI resources, and Department of Health and Human Services support of trans-agency public and private partnerships. By working together, the partners can accelerate their existing, independent efforts to improve our national capacity for high-quality, population-based U.S. cancer incidence data.

**Figure 6.** Surveillance Activities of National Coordinating Council for Cancer Surveillance (NCCCS) Member Organizations

Characteristic	Organization						
	American Cancer Society (ACS)	American College of Surgeons (ACoS) Commission on Cancer (COC)	National Cancer Institute (NCI)	Centers for Disease Control and Prevention (CDC)		National Cancer Registrars Association (NCRA)	North American Association of Central Cancer Registries (NAACCR)
				National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP)	National Center for Health Statistics (NCHS)		
Type of Organization	Voluntary	Professional	Government	Government	Government	Professional	Professional
Primary Cancer Surveillance Functions	Communication, education, and research; various collaborations with the ACoS	Education, accreditation, and support to clinicians and hospital registrars; patient care evaluation	Case registration; followup and survival; research	Case registration; cancer control research; state and school-based health surveys	Death registration/vital statistics; national health surveys	Education and support to registrars	Education and support to population-based registries
Data Program	NA	National Cancer Data Base (NCDB)	Surveillance, Epidemiology, and End Results (SEER) Program	National Program of Cancer Registries (NPCR); Behavioral Risk Factor Surveillance System (BRFSS); Youth Risk Behavior Survey (YRBS)	Vital Statistics; National Health Interview Survey (NHIS); National Health Examination Survey (NHANES); others	NA	Cancer in North America (CINA)
Sources of Data	Population-based surveys and cancer registries; NCHS	Hospital cancer registries	Population-based cancer registries	Population-based cancer registries; surveys	Death certificates; surveys	NA	Population-based cancer registries
Types of Data Reported	Annual forecasts of new cancer cases and deaths; descriptive statistics on cancer patient care and outcomes	Descriptive statistics regarding cancer patient care and outcomes	Incidence rates, mortality rates, survival, prevalence estimates, quality indicators	Incidence rates, number of cancer cases, death rates, prevalence estimates, health practices and behaviors	Death rates, number of cancer deaths, health practices and behaviors	NA	Incidence rates, number of cancer cases, mortality rates, number of deaths, data quality indicators
Publications	- Cancer Facts & Figures - Cancer Risk Report - Cancer - CA: A Journal for Clinicians	- Annual Review of Patient Care	- SEER: Annual Cancer Statistics Review - Journal of the National Cancer Institute - SEER Monographs	- Morbidity and Mortality Weekly Report	- National Vital Statistics Report - Vital Statistics of the U.S. - Healthy People 2000 Review - Health U.S. - Vital and Health Statistics	- Journal of Registry Management	- Cancer in North America

## Priority Area 3:

### Produce and disseminate a national report card on the cancer burden.

The SIG identified the following two needs within this priority area:

- A national report card on cancer is needed to report our Nation's progress in the effort to conquer cancer and to identify opportunities for further reduction in the cancer burden.
- Better methods are needed to disseminate the findings of the expanded surveillance research program and to communicate the availability of various resources (e.g., software programs, models, databases) developed by the CSRP that can be used by the extramural research community.

#### Research Opportunity 8

**Collect, analyze, and disseminate data on important cancer outcomes and trends in risk factor and screening behaviors as well as explanations for these trends in a National Cancer Report Card.** (*The cost for this effort is expected to be low; work should be initiated this year.*)

A National Report Card is needed to evaluate the Nation's performance in its ongoing efforts to confront the increasing burden of cancer on the population. The Report Card needs to include traditional cancer statistics on the burden from major cancers and all cancers together as well as trends in these rates. Such a Report Card also needs to capture the results of analytic research produced by the expanded CSRP as well as research results from other NCI programs, other federal agencies, and the extramural community that explain variations in the cancer burden as well as trends in the factors affecting it. With the expansion of NCI's surveillance program as outlined in this Implementation Plan, it is expected that such surveillance research and analysis will include the use of additional data on cancer patients, patterns of their care, variations in risk factors and screening data, and modeling. The Report Card should be published in a format accessible to the general public, the media, scientists, health organizations, legislators, and policymakers. There is a distinct challenge in addressing such a wide audience and avoiding an evaluation that is,

on the one hand, superficial and, on the other, esoteric. The NCI Bypass Budget is a good model for a public document of this type and, like the Bypass Budget, the Report Card should be accessible on the Internet.

The Report Card will evaluate our Nation's progress in characterizing and reducing the cancer burden. To assess this progress, it will be necessary to combine the resources and leadership of the NCI with the considerable contributions of other federal agencies, voluntary organizations, state and local health departments, advocates, and private citizens. The Report Card should be designed to tell the Nation how we are doing in this mammoth enterprise, but it also should be forward looking—identifying opportunities for further reduction in the cancer burden. The Report Card should serve the following purposes:

- For the lay public, it will present salient, understandable measures of the cancer burden and progress being made by research in reducing this burden.
- For the scientific community, it will contain critical cancer outcome measures and distilled interpretations of analytic work by the NCI, other federal agencies, and extramural scientists on the explanations for observed trends.
- For Congress and policymakers, it will demonstrate in understandable and “quotable” terms the benefits to the public of allocations of tax dollars to cancer research.
- For the NCI and other leaders in cancer control, it will provide a concise document to disseminate on the progress of cancer control that also highlights areas in need of future research.

The Report Card must capture cancer statistics trends in simple, comprehensible illustrations and provide easily understandable reasons for these changes despite the likelihood that the analysis underlying these reasons is both complex and substantial. Links to published scientific reports that provide details to interested audiences should be provided.

The SIG proposed a framework for the Report Card that links its elements to the cancer-related goals described in Healthy People 2010 as they relate to cancer-related risk factors, behaviors, incidence, survival, and mortality. These are national health promotion and disease prevention objectives produced and published by the Public Health Service. However, there are over 55 such goals when one considers those directly related to cancer as well as those related to tobacco, nutrition, and the environment that also are relevant. By selecting only the most salient measures, the Healthy People 2010 goals would be a starting point for developing the content and scope of the Report Card. The SIG expects that new measures will be identified in the process of developing the Report Card and that these measures will be included if sufficient data exist to support their inclusion.

All cancer sites need not be included in every Report Card. The SIG proposes that each Report Card cover the most common cancers (i.e., lung, breast, colon, and prostate cancers), and then focus selectively on those other cancer sites that reflect new trends or new knowledge. In addition, sites that represent opportunities for some positive action or behavior should be highlighted. The SIG recommends that the following measures of the cancer burden be included in the Report Card as they become available: incidence, stage-specific survival, mortality, tobacco, diet, alcohol, weight, physical activity, measures of inherited susceptibility, sun exposure, screening adherence, treatment, and environmental exposures.

Most of the content for the Report Card will be formulated by NCI staff who will review the issues, analyze the data, and prepare the various sections of the Report Card. Some content may come from other federal agencies, professional organizations, and extramural investigator-initiated research into relevant surveillance methodology and analysis. Contractual assistance will be required to design the layout, prepare the graphics, perform the desktop publishing, and print and disseminate the Report Card.

### Research Opportunity 9

**Develop a strategy for improved dissemination of information on the cancer burden via the Report Card and other NCI communications.** *(The cost of this effort is expected to be low to moderate; work should be initiated this year.)*

Better methods to disseminate the findings of the CSRP through other NCI communications beyond the Report Card are needed. These methods should ensure that more detailed data are made available to NCI and extramural investigators to facilitate the generation of hypotheses for epidemiologic studies, behavioral interventions, and public health programs. The data also should be available for studies to better explain the nature of the cancer burden and trends over time.

Part of the mission of the NCI is to disseminate statistical and related information to as wide an audience as possible. The CSRP currently disseminates information in reports (e.g., SEER Cancer Statistics Review), through journal articles, on the Internet (e.g., the SEER Web Site and the ARB Web Site), and through interaction with data users. Public-use and tailored datafiles as well as software tools are distributed by the CSRP free of charge upon request. The CSRP also collaborates with other government agencies and nongovernmental organizations to disseminate surveillance information. Two recent collaborative publications include a review of cancer incidence and mortality trends and a review of progress in improving diet to reduce cancer risk. Both of these publications have been quoted widely across agencies and in the press. SEER managers and Principal Investigators meet annually to exchange data and the CSRP makes updated cancer statistics available each year in both written and electronic formats. A more recent innovation is the establishment of "Stat Chats," where experts across the NCI and other NIH Institutes examine the latest cancer statistics and explore potential reasons for changes in trends. In addition, CSRP staff are actively involved with a number of surveillance-related organizations, including the NCCCS, the NAACCR, the International Agency for Research on Cancer (IARC), and CDC's NPCR. The CSRP should continue to support these information dissemination mechanisms.

With the expansion of the CSRP, there is a critical need for data management systems that properly store, organize, verify, analyze, present, and disseminate timely information. Such systems should provide users ready access to data at the national and local levels. Despite the huge collection of data for this purpose, there will be constant calls for more information in as yet unreported topic areas. The SIG recommends that

the CSRP continue to develop an understanding of the needs and perspectives of its data users and recognize emerging needs and potential areas for expanding its dissemination efforts. Specifically, the CSRP should develop an information dissemination strategy that stresses:

- Increased communication with current and potential data users.
- Development of methods to obtain feedback from users concerning the types of data needed and the desired formats.
- Development of a framework for reviewing and implementing user recommendations.
- Continued development of state-of-the-art computer software for data analysis, management, and dissemination.
- Expanded use of the Internet as an efficient and cost-effective mechanism to disseminate data.
- Increased data and resource utilization through improved packaging and marketing of the surveillance information and resources developed by the CSRP.

Also, cognitive research is needed to understand and aid the presentation of information to maximize understanding by the health research and care provider communities as well as the general public.

## Priority Area 4: Support molecular and genetics research for surveillance.

Within this priority area, the SIG identified the following two fundamental needs:

- Standardized, validated instruments for assessing family history of cancer are needed to track the population-based prevalence of family history of cancer.
- Research is needed to strengthen the surveillance of cancer-related genetic and molecular biomarkers.

### Research Opportunity 10

**Develop valid tools to assess family history of cancer and collection of data on the population prevalence of familial cancers.** (*The cost of this effort is expected to be moderate; work should be initiated this year.*)

There is a critical need for population-based prevalence data on cancer susceptibility genes, somatic mutations, and related outcome measures and risk factors, including family history of cancer. These data are needed to assess the population prevalence of cancer susceptibility genes, phenotypically identifiable rare syndromes such as heritable nonpolyposis colon cancer (HNPCC), and somatic alterations among people with specific cancers. In addition, data are needed to determine the population prevalence of environmental and behavioral risk factors for cancer in individuals who have a family

history of cancer or specific cancer susceptibility genes, as compared with the general population. The impact of family history and cancer susceptibility genes on population-based cancer incidence, mortality, and other outcome measures also should be examined as well as the implications that they have for screening high-risk individuals for genetic alterations and for cancer. Research is needed to assess the impact of cancer susceptibility genes and somatic tumor alterations on prognosis and treatment of cancer patients in the general population.

Tracking the population-based prevalence of family history of cancer, including known major inherited cancer syndromes, will provide critical data for both cancer research and public health planning. There is a need to develop standardized, validated instruments for assessing family history of cancer as part of the surveillance program as well as procedures for assuring the confidentiality of the individual-level data that are collected.

The SIG proposes that the CSRP develop a computer-based questionnaire to ascertain family structure and family history of cancer in first- and second-degree relatives, which will be pretested and validated in various population subgroups. The instrument should be distributed to the general extramural community and also used in a large nationally representative surveillance study to estimate the number of individuals with inherited can-

cer syndromes (e.g., HNPCC; breast/ovarian). It also should allow for the stratification of the general population by levels of relevant behavioral and environmental risk factors. This research could be supported through one or more mechanisms such as SEER Special Studies, a contract mechanism, an interagency agreement, or a cooperative agreement. The national survey of prevalence requires the development of a national probability sample and might best be accomplished through a contract or interagency agreement mechanism. However, other mechanisms such as PAs or RFAs might be used to support the development of estimates for local and other defined populations.

### Research Opportunity 11

**Investigate the feasibility of expanding population-based molecular and genetic biomarker studies within the Cancer Surveillance Research Program.** *(The cost for this effort is expected to be moderate to high; work should be initiated within the next 1-2 years.)*

The traditional role of the SEER Program in generating hypotheses for etiologic studies and for providing a source of cases for case-control studies would be enhanced by the collection of DNA-containing biospecimens on newly diagnosed cancer patients. To further enhance the value of the CSRP

for such studies, methods for making rapid case ascertainment capabilities more broadly available should be explored. This would be of particular value for case-control studies of less common neoplasms where early mortality is likely.

Furthermore, the surveillance of cancer-related genetic and molecular biomarkers would be strengthened by developing criteria for selecting specific biomarkers, enhancing infrastructure for collecting and archiving biospecimens, and addressing the ethical, legal, and social issues associated with these activities.

The SIG recommends that the CSRP initiate methodologic and pilot studies to evaluate the feasibility of biomarker studies within SEER and other components of the CSRP. In support of these studies, the CSRP should assess available biotechnology for accurate measurement of biomarkers, identify suitable study populations, and enhance current infrastructures for rapid case ascertainment and data tracking systems. Because collection and management of data and biospecimens differ among SEER sites and other networks, the CSRP should promote efforts to pool and standardize data for centralized use and work closely with existing structures—such as the Cancer Genetics Network and the Cancer Family Registries—to ensure careful coordination of efforts.

## Priority Area 5: Develop a training strategy for cancer surveillance research.

### Research Opportunity 12

**Identify specific training needs related to surveillance sciences and develop a plan to incorporate surveillance into mechanisms for training cancer prevention and control scientists.** *(The cost for this effort is expected to be moderate; work should be initiated with the next 3-5 years.)*

NCI's cancer training program currently supports a broad range of training activities in the finest institutions in the country as well as individual fellowships, career awards, and education grants. Because those trained in cancer research today will form the intellectual foundation for basic, clinical, and population sciences 10 years from now, training activities

sponsored by the NCI must continually anticipate the human resource needs of cancer research in the future. According to the Bypass Budget for Fiscal Year 2000, the NCI will meet this challenge by pursuing four independent training and education strategies:

- Maintaining the critical mass of independent and basic scientists studying cancer at the most fundamental levels of genetics and molecular biology.
- Encouraging a greater proportion of well-trained basic scientists currently engaged in research on model systems to develop interests in model systems for human biology and human disease.

- Attracting more young physicians, other health care professionals, and public health specialists into cancer research; of particular importance will be continuing programs that will develop a larger contingent of physicians, other health care professionals, and public health specialists in biostatistics, epidemiologic, behavioral, and other prevention and control sciences.
- Using education grants to improve the curricula for health care and public health students, and improving community education and information dissemination programs.

The SIG noted that no formal training in surveillance research currently exists. With the growing demand for collecting additional data and the increasing complexity of analyzing such data, there is a need for training programs in surveillance sciences. The SIG recommends that specific training needs related to cancer surveillance be identified. For example, training is needed to help epidemiolo-

gists comprehend the potential usefulness of GIS in cancer surveillance and how to use geospatial information.

The CSRP should play a significant role in working with members of the NCCCS, professional organizations, and NCI's training and fellowship programs to develop and implement a plan to address these training needs. The CSRP also should assess current and future personnel needs. The plan should address short-term intensive training opportunities for established scientists to broaden their information base in the surveillance sciences as well as incorporating surveillance sciences into the pre- and postdoctoral training mechanisms funded by the member organizations of the NCCCS. The NCI could use existing and new award mechanisms to educate modelers, biostatisticians, and economists as well as health services, prevention, control, population, and outcomes research scientists in cancer surveillance sciences.

**Cancer Surveillance Data and Research Resources are Available on the Web**

**SEER Web Site:**

<http://www-seer.ims.nci.nih.gov/>

**ARB Web Site:**

<http://www-dccps.ims.nci.nih.gov/ARB/>

# Appendix A: The NCI Surveillance Implementation Group

## Co-Chairs:

**Robert A. Hiatt, M.D., Ph.D.**

Deputy Director  
Division of Cancer Control and Population Sciences  
National Cancer Institute  
Bethesda, MD

**Nicole Urban, Sc.D.**

Associate Member  
Cancer Prevention Research Program  
Division of Public Health Sciences  
Fred Hutchinson Cancer Research Center  
Seattle, WA

## Members:

**Rachel Ballard-Barbash, M.D., M.P.H.**

Chief, Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Marsha G. Davenport, M.D., M.P.H.**

Chief Medical Officer  
Office of Strategic Planning  
Health Care Financing Administration  
Baltimore, MD

**Nancy Breen, Ph.D.**

Economist  
Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Robert W. Day, M.D., M.P.H., Ph.D.**

President and Director Emeritus  
Member, Division of Public Health Sciences  
Fred Hutchinson Cancer Research Center  
Seattle, WA

**Martin Brown, Ph.D.**

Chief, Health Services and Economics Section  
Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Manuel de la Puente, Ph.D.**

Chief, Ethnic and Hispanic Branch  
U.S. Bureau of Census  
Population Division  
Washington, DC

**David M. Burns, M.D.**

University of California, San Diego  
Tobacco Control Policies Project  
San Diego, CA

**David DeMets, Ph.D.**

Professor of Statistics and Biostatistics  
Chair, Dept. of Biostatistics and Medical Informatics  
University of Wisconsin  
Madison, WI

**Vivien W. Chen, Ph.D.**

Director/Epidemiologist  
Louisiana State University Medical Center  
Department of Pathology  
New Orleans, LA

**Susan Devesa, Ph.D.**

Chief, Descriptive Studies Section  
Biostatistics Branch, Epidemiology and Biostatistics  
Program  
Division of Cancer Epidemiology and Genetics  
National Cancer Institute  
Bethesda, MD

**Brenda K. Edwards, Ph.D.**

Associate Director  
Cancer Surveillance Research Program, DCCPS  
National Cancer Institute  
Bethesda, MD

**Virginia L. Ernster, Ph.D.**

Professor and Vice Chair  
Department of Epidemiology and Biostatistics  
School of Medicine  
University of California at San Francisco  
San Francisco, CA

**Ellen Feigal, M.D.**

Deputy Director  
Division of Cancer Treatment and Diagnosis  
National Cancer Institute  
Bethesda, MD

**Jack J. Feldman, Ph.D.**

Center for Health Affairs, Project Hope  
(Former Deputy Director, NCHS)  
Chevy Chase, MD

**Eric [Rocky] Feuer, Ph.D.**

Chief, Surveillance Modeling and Methods Section  
Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Andrew N. Freedman, M.S., Ph.D.**

Epidemiologist  
Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Barry Graubard, Ph.D.**

Biostatistics Branch  
Division of Cancer Epidemiology and Genetics  
National Cancer Institute  
Bethesda, MD

**Robert M. Groves, Ph.D.**

Director, Joint Program in Survey Methodology  
University of Maryland  
College Park, MD

**James W. Hampton, M.D.**

Medical Director  
Troy and Dollie Smith Cancer Center  
Oklahoma City, OK

**Benjamin F. Hankey, Sc.D.**

Chief, Cancer Statistics Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Linda C. Harlan, Ph.D.**

Epidemiologist  
Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Robert N. Hoover, M.D., Sc.D.**

Director  
Epidemiology & Biostatistics Program Division of  
Cancer Epidemiology and Genetics  
National Cancer Institute  
Bethesda, MD

**Larry Kessler, Sc.D.**

Director, Office of Surveillance and Biometrics  
Center for Device and Radiological Health  
Food and Drug Administration  
Rockville, MD

**Sue Krebs-Smith, Ph.D.**

Chief, Risk Factor Monitoring and Methods Section  
Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Shiriki Kumanyika, Ph.D., M.P.H., R.D.**

Professor, Department of Human Nutrition and  
Dietetics  
University of Illinois at Chicago  
Chicago, IL

**Charles E. Lynch, M.D., Ph.D.**

Associate Professor  
University of Iowa, College of Medicine  
Department of Preventive Medicine  
Iowa City, IA

**Mark B. McClellan, M.D., Ph.D.**

Deputy Assistant Secretary for Economic Policy  
Department of the Treasury  
Washington, DC

**Nancy E. Mueller, Sc.D.**

Professor of Epidemiology  
Harvard School of Public Health  
Boston, MA

**Christopher J. L. Murray, M.D., Dr.P.H.**  
Associate Professor of International  
Health Economics  
Harvard School of Public Health  
Center for Population and Development Studies  
Cambridge, MA

**G. Iris Obrams, M.D., Ph.D.**  
Associate Director  
Epidemiology and Genetics Research Program,  
DCCPS  
National Cancer Institute  
Bethesda, MD

**Edward B. Perrin, Ph.D.**  
Professor, Department of Health Services  
University of Washington  
Seattle, WA

**Janice Platner, J.D.**  
Massachusetts Cancer Registry  
Advisory Committee  
Somerville, MA

**Lynn A. G. Ries, M.S.**  
Health Statistician  
Cancer Statistics Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

## Ad Hoc Members:

**Joseph F. Fraumeni, Jr., M.D.**  
Director  
Division of Cancer Epidemiology and Genetics  
National Cancer Institute  
Bethesda, MD

**Barbara K. Rimer, Dr.P.H.**  
Director  
Division of Cancer Control and Population Sciences  
National Cancer Institute  
Bethesda, MD

**Gerard Rushton, Ph.D.**  
Professor  
Department of Geography  
The University of Iowa  
Iowa City, IA

**Judith Swan, M.H.S.**  
Public Health Advisor  
Cancer Surveillance Research Program, DCCPS  
National Cancer Institute  
Bethesda, MD

**Lillian Tom-Orme, Ph.D., M.P.H., R.N.**  
University of Utah  
Huntsman Cancer Institute  
Division of Public Health Sciences  
Salt Lake City, UT

**Louise Wideroff, M.S.P.H., Ph.D.**  
Epidemiologist  
Applied Research Branch, CSRP, DCCPS  
National Cancer Institute  
Bethesda, MD

**Robert E. Wittes, M.D.**  
Director  
Division of Cancer Treatment and Diagnosis  
National Cancer Institute  
Bethesda, MD



# B

## Appendix B: Cancer Sur in the United States

There is no single national cancer surveillance system in the United States that collects data on all cases diagnosed each year. Distinct segments of the U.S. population are covered by separate programs sponsored by government or private organizations, and these programs provide reliable data to their varying audiences. All the programs are built on a foundation of cancer registries that exist or are being established throughout the country to record and report on cases in their service areas. Although significant steps have been taken to move toward the use of a uniform data set, differences in data collection, analysis, and reporting place an extra burden on registries that report to more than one surveillance program. In addition, the inconsistencies of data sets present obstacles to their compilation for collaborative use. The various organizations involved in cancer surveillance are described below.

### The National Coordinating Council for Cancer Surveillance (NCCCS)

The NCCCS was created to provide a forum for examining the current state of cancer surveillance operations, to identify the broad issues involved, and to recommend practical approaches that will facilitate the work of registries and contribute to the goal of coordinating data collection and improving data quality across the Nation. Its mission is to coordinate cancer surveillance activities within the United States through communication and collaboration among major national cancer organizations, ensuring that the needs of cancer patients and the communities in which they live are fully served, that scarce resources are maximally used, and that the burden of cancer in the United States is adequately measured and ultimately reduced.

Although the NCCCS does not have direct authority to implement recommendations, its membership represents the major cancer surveillance programs in place today. The key issues that the NCCCS will address affect all of these programs, such as the demands created by changing uses of cancer data, which in turn increase the resources necessary to support registry systems. With potential users ranging from clinicians and statisticians to hospital administrators or patients and their families, coordinated decisions on formats and retrieval systems are essential. As new or different uses of the data develop, maintenance of data quality will be an important concern integral to each organization involved in cancer registration.

In defining cancer surveillance for the purpose of national coordination, the NCCCS currently has restricted its interpretation to the measurement of incidence, mortality, morbidity, and survival. A more global view of surveillance would include diverse measures of cancer risk and cancer management in clinical, family, or community settings, lifestyle factors, screening utilization, behavioral influences, genetic predisposition, or environmental exposures. In 1990, a special committee convened by the National Cancer Institute (NCI) at the request of Congress, reported on the relevance of socioeconomic, cultural, and health system factors to assessing progress against cancer in the United States. The NCCCS, in voting to confine its opening efforts to cancer case registration, acknowledged the importance of these other factors in the overall concept of cancer surveillance. However, the initial focus of the NCCCS will be the more basic foundation of data upon which much of cancer control planning, implementation, and evaluation is based.

## Scope of the Cancer Registration Effort

Collection of cancer data in the United States has developed under several systems of registries, and the data collected vary according to the purposes served. Generally, they include hospital registries, which may be part of a facility's cancer program, and population-based registries, which are usually associated with state health departments or institutions to which they delegate authority. Hospital registries provide complex data for the evaluation of care within the hospital, and they are the primary source of data for state registries. Population-based registries, such as those associated with state health departments, record and consolidate information regarding all cases diagnosed within a specific geographic area and therefore provide data that can determine rates across regions of the country.

Cancer registration is primarily performed by cancer registrars in hospital-based and population-based registries. These registrars track down and locate the wide variety of cancer data that is required to be collected. Whatever the venue, registries function best when cancer registrars have met stringent standards of training, testing, and continuing education—the highest level being Certified Tumor Registrars (CTR). The experiences of cancer registrars have led to current efforts to coordinate the various programs to prevent redundancies and create a nationally operational system of cancer surveillance. One of the compelling reasons for a national system is that the process of data collection and reporting is time consuming and labor intensive. Moreover, most hospital-based registrars code and report data to more than one surveillance program.

Cancer registration begins with “casefinding,” or identification of individuals with cancer who have sought care at hospitals and other medical care settings. Most often, the patient's physician initiates the data record by noting in the medical record the cancer site and type, patient demographics, and extent of disease or stage. Some surveillance programs require that the registrar abstract additional information from the patient record such as type of treatment and annual followup for disease recurrence and survival. The sources of data include patient medical records, laboratory records, appointment logs, and administrative or billing records.

The information, often abstracted in considerable detail, is coded, checked for accuracy, and sent forward to a central registry database. Depending on the target database, the codes may vary significantly, and cancer registrars usually are required to apply more than one coding system. A hospital registry traditionally incorporates all of this information for evaluating cancer care. Another common use for such data is in studies that compare patterns of care among providers, population subsets, or geographic regions. Other applications of cancer registry data have developed in recent years, such as linking patients to available research protocols and collecting information on biological factors for diagnosis or staging.

Population-based registries in metropolitan areas and states collect and consolidate information from multiple reporting facilities, which can include hospitals, physicians' offices, nursing homes, pathology laboratories, ambulatory care facilities, radiation and chemotherapy treatment centers, and other cancer care arenas within defined geographic areas. Frequently established by state legislative mandate, these registries can use area census data to calculate incidence rates essential for the evaluation of cancer control efforts, the investigation of cancer clusters, and comparisons of patient care. Registries can link with other information sources such as state and federal vital records, voter and motor vehicle registration databases, or Medicare files for the purpose of followup and analysis of survival. Both hospital- and population-based registrars may take a more active role by directly contacting the patients or their cancer care providers for followup information.

There currently are three major cancer surveillance programs in the United States—the National Cancer Data Base; the Surveillance, Epidemiology, and End Results Program; and the National Program of Cancer Registries.

### National Cancer Data Base

Established in 1989, the National Cancer Data Base (NCDB) is a program of hospital and selected ambulatory care registries of the Commission on Cancer (COC), which is administered by the American College of Surgeons (ACoS). It is jointly sponsored by the ACoS and the American Cancer Society (ACS), for the purpose of ensuring quality cancer care by providing data for evaluation of pa-

tient management within hospitals and other treatment centers and for comparisons between institutions or regions of the country.

Approximately 1,500 cancer treatment centers in the United States contribute to the NCDB, and in 1996 this became a requirement for COC-approved cancer programs. The NCDB prepares annual reports for individual institutions that contain tables describing the reporting facility's cancer activities and comparisons with aggregate data. The *NCDB Annual Review of Patient Care* provides nationwide data on trends and patterns of care for specific cancer sites. In addition, the NCDB analysts prepare many special reports upon request to aid the ACS area divisions and educators in pinpointing locations where increased cancer control efforts are needed. University researchers also are able to use the analytic file for specific studies of interest. In addition, subcommittees of specialists design and monitor patient care evaluation (PCE) studies to provide timely information on patterns of care related to geographic, socioeconomic, and clinical factors. To disseminate results of PCE studies, the data are presented at professional meetings and reports are submitted for publication in peer-reviewed journals.

## Surveillance, Epidemiology, and End Results Program

The Surveillance, Epidemiology and End Results (SEER) Program is a system of population-based registries administered by the NCI. SEER was established in 1973 to provide continuous coverage in certain regions of the United States with authorizing legislation in place for central data collection. The NCI contracts with nonprofit organizations to collect data on all new cancer cases diagnosed in their geographic locations. Cases are followed up annually to determine survival. These data, along with data on cancer-related deaths from the National Center for Health Statistics (NCHS), are analyzed to provide incidence, mortality, and survival rates.

The SEER Program collects and publishes cancer incidence and survival data from 11 population-based cancer registries and 2 supplemental registries covering approximately 14 percent of the U.S. population. The 11 SEER registries are located in five states (Connecticut, Iowa, New Mexico, Utah,

and Hawaii) and six metropolitan areas (Detroit, San Francisco/Oakland, Seattle/Puget Sound, San Jose/Monterey, Atlanta, and Los Angeles). The two supplemental registries include American Indians in Arizona, which are registered by the New Mexico SEER registry, and 10 rural, predominantly black counties in Georgia, which are registered by the Atlanta registry. The supplemental registries were added to increase coverage of minority populations. Currently, 25 percent of the American Hispanic population, 41 percent of the Asian/Pacific Islanders population (43 percent of all Chinese Americans and 60 percent of all Japanese Americans), 27 percent of the American Indian and Alaska Native populations, and 12 percent of the African American population reside in SEER areas (see Tables 1 and 2). SEER is reasonably representative of the U.S. population for purposes of cancer surveillance based on analyses of mortality trends in SEER areas compared to the total United States.

Quality assurance has been a top priority for the SEER Program since its inception, with onsite monitoring, data editing, casefinding audits, and reabstracting of cases as part of Program activity as well as extensive educational workshops and instruction manuals. SEER data standards for reporting have established a case ascertainment of 98 percent and a followup rate of 95 percent for all ages combined. Recently, SEER personnel have developed software for data analysis and a public-use electronic file available through the Internet. Uses of SEER data include research on cancer trends and their relationship to cancer control efforts such as screening programs or passage of tobacco-related legislation; identification of populations at risk for higher rates of cancer; comparisons of cancer incidence, mortality, and risk factors in geographic areas; and studies of patterns of care.

## National Program of Cancer Registries

Complete and timely collection of population-based data requires a commitment of personnel and money as well as legal authority. The National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention (CDC) supports population-based registries in state health departments. It was authorized by the Cancer Registries Amendment Act in 1992, in response to the obser-

vation that, although many states had established registries, only the SEER states and California covered 95 percent or more of their respective populations. In 1990, 10 of the 50 states had no central registry. A number of state and federally sponsored cancer control programs carried out in the 1980s and early 1990s, including NCI's Data-Based Intervention Research (DBIR) Program, pointed to the need for state cancer registries to identify and analyze data that would be relevant to developing, implementing, and evaluating cancer control programs and plans to meet state and local needs.

Congress has appropriated resources to the CDC to develop model state legislation and regulations that would require hospitals and health care practitioners that diagnose or treat cancer to report all cases and ensure access to medical records for state registry personnel. The CDC was authorized to set and monitor national standards for data completeness, timeliness, and quality. By the end of Fiscal Year 1998, the NPCR was providing support to 45 states and the District of Columbia, Puerto Rico, the Virgin Islands, and Palau. Of these, 13 states were developing new registries and 36 were in the process of enhancing existing registries. When fully operational, the NPCR will collect cancer incidence data on 97 percent of the U.S. population. Tables 1 and 2 provide data on the racial/ethnic coverage of the NPCR.

The quality standards of the North American Association of Central Cancer Registries (NAACCR), a cancer registry coordinating and oversight organization, are being applied by the CDC to the NPCR-funded registries along with audits, onsite monitoring, use of standardized software, and educational

programs. Uses of state registry data will include health planning and resource allocation, evaluation of cancer control programs, identification of populations at risk, and comparisons of cancer incidence across specific geographic areas. The state registries also will serve as population-based sampling frames for epidemiologic and clinical research.

## Communicating Cancer Surveillance Data

The ACoS disseminates information from the NCDB through the *Annual Review of Patient Care* publication. The NCI distributes SEER data through the *SEER: Annual Cancer Statistics Review*, various monographs, the *Journal of the National Cancer Institute*, and the SEER Web Site. The CDC disseminates surveillance data through a variety of publications, including the *Morbidity and Mortality Weekly Report*, the *National Vital Statistics Report*, and the *Vital Statistics of the U.S.* publication. A number of other organizations use and report surveillance data resulting from the NCDB, SEER, and/or NPCR. For example, the ACS publishes *Cancer Facts & Figures* and the *Cancer Risk Report* as well as the journals *Cancer* and *CA: A Journal for Clinicians*. The NAACCR publishes the monograph entitled *Cancer in North America (CINA)* each year. The most recent CINA reported cancer incidence and mortality data from 40 central cancer registries for the years 1990 to 1994. Data from 19 of these cancer registries, which cover about 38 percent of the U.S. population, were used to compute the U.S. combined cancer incidence rates (see Tables 1 and 2 for more data on the racial/ethnic coverage of CINA).

**Table 1.**

Proportion of U.S. Racial/Ethnic Groups Residing in SEER Registries, the 1990-1994 CINA Poolable Registries, and the NPCR Program Registries

Race/Ethnicity	SEER (11 Registries)	SEER (11 Registries) and SEER Supplemental	CINA (19 Poolable Registries)	NPCR Enhancement Programs	NPCR Planning Programs	Total United States
White	12.48%	12.50%	37.16%	58.49%	10.23%	100%
Black	12.07%	12.24%	31.32%	60.09%	14.82%	100%
American Indian, Eskimo, and Aleut	16.41%	26.55%	37.17%	32.93%	25.85%	100%
Chinese	43.53%	43.53%	61.25%	39.82%	2.33%	100%
Filipino	49.09%	49.10%	78.54%	23.87%	2.36%	100%
Japanese	59.79%	59.79%	80.07%	19.54%	3.04%	100%
Vietnamese	31.19%	31.19%	62.84%	35.63%	4.56%	100%
Korean	33.69%	33.69%	58.59%	44.50%	3.93%	100%
Other Race	28.20%	28.20%	56.93%	44.72%	1.77%	100%
All Races	13.89%	14.01%	38.06%	57.28%	10.36%	100%
Hispanic, All Races	24.99%	24.99%	51.27%	49.08%	1.57%	100%

**Table 2.**

Comparison of Racial/Ethnic Distribution of the United States and SEER Registries, the 1990-1994 CINA Poolable Registries, and the NPCR Programs

Race/Ethnicity	SEER (11 Registries)	SEER (11 Registries) and SEER Supplemental	CINA (19 Poolable Registries)	NPCR Enhancement Programs	NPCR Planning Programs	Total United States
White	72.19%	71.70%	78.46%	82.05%	79.36%	80.35%
Black	10.46%	10.51%	9.90%	12.63%	17.21%	12.03%
American Indian, Eskimo, and Aleut	0.96%	1.54%	0.79%	0.47%	2.02%	0.81%
Chinese	2.08%	2.06%	1.07%	0.46%	0.15%	0.66%
Filipino	2.02%	2.00%	1.18%	0.24%	0.13%	0.57%
Japanese	1.50%	1.49%	0.73%	0.12%	0.10%	0.35%
Vietnamese	0.54%	0.53%	0.39%	0.15%	0.10%	0.24%
Korean	0.78%	0.77%	0.49%	0.25%	0.12%	0.32%
Other Race	9.48%	9.40%	6.98%	3.64%	0.80%	4.67%
All Races	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Hispanic, All Races	15.85%	15.71%	11.86%	7.55%	1.33%	8.81%





# **Appendix C: Major Surveys Databases Used by NCI's Cancer Prevention Research Program**

Survey/Database and Sponsoring Organization	Years Covered	Phase of Cancer: Prevention—Surveillance of Risk Factors	Brief Description
First National Health and Nutrition Examination Survey (NHANES I) — National Center for Health Statistics (NCHS)	1971-1974		Civilian, noninstitutionalized population of the conterminous United States, 1 to 74 years of age. Dietary intake (one 24-hour recall), food frequency, socioeconomic and demographic information, biochemical analyses of blood and urine, physical examination, and body measurements. Sample includes nearly 21,000 examined persons.
Nationwide Food Consumption Survey (NFCS) — U.S. Department of Agriculture (USDA)	1977-1978		Private households in the 48 conterminous states and the individuals in those households (all income and low income). For households: quantity (pounds), money value (dollars), and nutritive value of food used. For individuals: 1-day and 3-day food and nutrient intakes by individuals of all ages, name and times of eating occasions, and sources of food obtained and eaten away from home. Data collected over 3 consecutive days using a 1-day recall and a 2-day record. Intakes are available for 15 nutrients and food components, from nearly 31,000 individuals.
Second National Health and Nutrition Examination Survey (NHANES II) — NCHS	1976-1980		Civilian, noninstitutionalized population of the United States; 6 months to 74 years of age. Dietary intake (one 24-hour recall), food frequency, socioeconomic and demographic information, biochemical analyses of blood and urine, physical examination, and body measurements. Sample includes over 25,000 persons interviewed and over 20,000 examined.
NHANES I Epidemiologic Followup Survey — NCHS	1982-1984, 1986		Personal interview for survivors and proxy interview for decedents and incapacitated persons including medical history, history of hospitalization, functional status, medication usage, smoking history, alcohol history, psychological status, food frequency, and physical activity; physical measurements of pulse, blood pressure, and weight; death certificates; and hospital and nursing home records for overnight stays. Over 12,000 persons were interviewed between 1982-84.
Nationwide Food Consumption Survey (NFCS) — USDA	1987-1988		Households in the 48 conterminous United States and individuals residing in those households. The survey was composed of two samples: a basic sample of all households and a low-income sample of households with incomes <130% of the poverty threshold. For households: quantity (pounds), money value (dollars), and nutrient value of food used. For individuals: 1-day and 3-day food and nutrient intakes by individuals of all ages, names and times of eating occasions, and sources of food obtained and eaten away from home. Data collected over 3 consecutive days using a 1-day recall and a 2-day record. Intakes are available for 28 nutrients and food components, from over 10,000 individuals.
Third National Health and Nutrition Examination Survey (NHANES III) — NCHS	1988-1994		Civilian, noninstitutionalized population 2 months of age and older. Over sampling of non-Hispanics, blacks, and Mexican Americans, children <6 years of age, and adults aged >60 years. Dietary intake (one 24-hour recall and food frequency), socioeconomic and demographic information, biochemical analyses of blood and urine, physical examination, body measurement, blood pressure measurements, bone densitometry, dietary and health behaviors, and health conditions. Two additional 24-hour recalls for participants 50 years of age and older. Sample includes about 34,000 interviewed persons and about 31,000 examined.
Continuing Survey of Food Intakes by Individuals (CSFII) — USDA	1989-1991		Individuals in households in the 48 conterminous States. The survey was composed of two separate samples: households with incomes at any level (basic sample) and households with income <130% of the poverty thresholds (low-income sample). One-day and 3-day food and nutrient intakes by individuals of all ages, names and times of eating occasions, and sources of food obtained and eaten away from home. Data collected over 3 consecutive days by use of a 1-day recall and a 2-day record. Intakes are available for 28 nutrients and food components, from about 15,000 individuals.

Survey/Database and Sponsoring Organization	Years Covered	Brief Description
<b>Phase of Cancer: Prevention—Surveillance of Risk Factors (continued)</b>		
5 A Day Baseline and Followup Surveys — National Cancer Institute (NCI)	1991, 1997	Telephone surveys of a randomly selected nationally representative sample of U.S. adults (sample size of about 2,800 in 1991 and about 2,600 in 1997). The purpose of the baseline survey was to assess Americans' fruit and vegetable intakes, and their awareness, knowledge, and attitudes about diet and nutrition. To allow assessment of trends similar questions were asked in the two surveys.
Continuing Survey of Food Intakes by Individuals (CSFI) — USDA	1994-1996	Individuals in households in the 48 conterminous States. The survey was composed of two separate samples: households with incomes at any level (basic sample) and households with income <130% of the poverty threshold (low-income sample). Two-day food and nutrient intakes by individuals of all ages, names and times of eating occasions, and sources of food obtained and eaten away from home. Data collected over 2 non-consecutive days by use of two 1-day recalls. Intakes are available for 28 nutrients and food components, from about 16,000 individuals.
National Food Supply Data — USDA	1909-present	Food supply data represent the quantities of foods and nutrients that disappear into the retail distribution channels each year. The data are derived by first summing together the total quantities of foods produced in this country, inventories remaining from the past year, and all imported food. Subtracted from this are current year-end inventories, exported food, and food used for non-food purposes. Includes data for approximately 400 different foods.
ADD Health (National Longitudinal Study of Adolescent Health) — National Institute of Child Health and Human Development (NICHD) with support from NCI and other NIH Institutes	1995-1996, with followup plans for 5-6 years later	ADD Health (National Longitudinal Study of Adolescent Health) is a U.S. national probability sample school-based study of the health-related behaviors of adolescents placed within a social context. NICHD funded this study with supplemental support from the NCI and other NIH Institutes. Information about alcohol, tobacco, diet, physical activity, ultraviolet exposure (both artificial and natural sunlight) important for cancer risk factor surveillance has been obtained. In 1995, 90,000 students in grades 7 through 12 (ages 12-17) attending 145 schools answered brief questionnaires about their lives, including their health, friendships, self-esteem, and expectations for the future. Then 20,000 in-home interviews of students and 18,000 parents (one parent per student, usually the mother) were conducted between April and December 1995. A followup of 15,000 adolescents interviewed again at home was conducted in 1996. In 1995 and 1996, school administrators completed a survey about school policies, teacher characteristics, health service information, and student body characteristics.
Youth Risk Behavior Survey (YRBS) — National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP)	1990, 1991, 1993, 1995, 1997	The YRBS monitors six categories of priority health-risk behaviors, including tobacco use, alcohol and drug use, sexual behaviors, unhealthy dietary behaviors, and physical inactivity. It includes both a national school-based survey of school students in grades 9 through 12 conducted by the CDC and state and local school-based surveys conducted by state and local education agencies. For the national survey, approximately 16,000 questionnaires are completed in about 150 schools.
Tobacco Use Supplement (TUS) for the Current Population Survey (CPS) — Census Bureau and NCI	1992-1993, 1995-1996, 1998-1999	The TUS/CPS collects data from the civilian noninstitutionalized population on tobacco use and smoking prevalence, smoking intervention dissemination of workplace smoking policies and cessation programs as well as medical and dental advice to stop smoking, and changes in smoking norms and attitudes. 294,000 total respondents interviewed in 1992-93 and 247,000 total respondents interviewed in 1995-96; a sample size of approximately 247,000 is projected for 1998-99.

Survey/Database and Sponsoring Organization	Years Covered	Brief Description
<b>Phase of Cancer: Prevention—Surveillance of Risk Factors (continued)</b>		
State Cancer Legislative Database (SCLD) — NCI	1989-1999	The SCLD maintains information about state legislation and regulation addressing cancer-related topics, including tobacco, breast cancer, genetics, cervical cancer, prostate cancer, testicular cancer, colon cancer, ovarian cancer, access to state-of-the-art cancer treatment, cancer registries, and selected issues related to diet and nutrition, employment discrimination, environmental exposure, and occupational exposure. The SCLD Program collaborated with CDC's Office on Smoking and Health to conduct an analysis of State Laws on Tobacco Control—United States, 1995. The Program publishes the SCLD Update, a quarterly newsletter that reviews all substantive state cancer-related legislation enacted and resolutions adopted each quarter. The SCLD Updates Index, which can be searched directly on the Internet, promotes public access to previously compiled summaries of laws that were reported in SCLD Update.
Behavioral Risk Factor Surveillance System (BRFSS) — NCCDPHP	1984-1998	The BRFSS is a state-based, random-digit-dialed telephone survey of the U.S. civilian, noninstitutionalized adult population (ages 18 years and older). Data are obtained from all 50 states and approximately 90,000 persons respond to the survey. The BRFSS collects information on behavioral risk factors that are related to one or more of the 10 leading causes of death. These factors include tobacco use, alcohol consumption, diet and health, weight, drug use, screening utilization, and health insurance status.
National Health Interview Survey, Cancer Epidemiology and Cancer Control Supplements — NCHS and NCI	1987, 1992, 2000	The third NHIS Cancer Control Supplement (previously fielded in 1987 and 1992) will be fielded in 2000 and will collect data from a national representative sample (sample size for adults 18 and over is 24,000 in 1987, 12,000 in 1992, and 48,000 planned in 2000) to allow tracking of trends in cancer-related risk factors, health behaviors including use of screening, and knowledge and attitudes regarding these factors. New additions to the 2000 NHIS CCS include questions related to knowledge, attitudes, and use of genetic testing, family history of cancer, and more questions related to cancer survivorship issues.
<b>Screening/Detection</b>		
National Health Interview Survey, Cancer Epidemiology and Cancer Control Supplements — NCHS and NCI	1987, 1992, 2000	See Prevention — Surveillance of Risk Factors
Behavioral Risk Factor Surveillance System (BRFSS) — NCCDPHP	1984-1998	See Prevention — Surveillance of Risk Factors
Breast Cancer Surveillance Consortium (BCSC) — NCI	1995-2000	In response to the Mammography Quality Standards Act (MQSA) mandate to fund such research, the NCI established the Breast Cancer Surveillance Consortium in 1994. Eight sites throughout the U.S. are participating in the BCSC. The major objectives of the Surveillance Consortium are to: enhance our understanding of the operation and conduct of breast cancer screening in the United States; foster collaborative research among Surveillance Consortium participants to assess factors associated with variation in mammography practice, accuracy, and subsequent diagnostic evaluation; and provide a foundation for the conduct of clinical and basic science research that can improve understanding of breast cancer etiology and prognosis. The Surveillance Consortium database includes patient demographic and health history, radiologic, diagnostic followup, and pathologic outcomes. By the year 2000, the database will contain information on nearly 3.2 million screening mammographic examinations and over 24,000 breast cancer cases.

Survey/Database and Sponsoring Organization	Years Covered	Brief Description
<b>Screening/Detection (continued)</b>		
National Survey of Mammography Facilities (NSMF) — NCI	1992	Questionnaires were mailed or presented by telephone to a nationally representative sample of facilities offering mammography services in the United States. Of the 1,165 facilities eligible for the survey, 1,057 submitted completed questionnaires. The NSMF collected information on the year of accreditation by the American College of Radiology (ARC), whether facilities distinguish between screening and diagnostic mammography, monthly volume, organizational affiliation, qualifications and training of key personnel, equipment, average examination charges, components of examination (whether breast physical examination and/or instruction in breast self examination (BSE) are routinely offered), and performance of procedures commonly associated with quality assurance such as annual surveys by radiologic physicists, daily sensitometry, peer review of film interpretation, and monitoring of repeat films due to technical problems.
Prostate Cancer Outcomes Study — NCI	Diagnosed 10/94 through 10/95	Study outcomes in patients newly diagnosed with prostate cancer. Focus on practice patterns and quality of life. Patient interviews at 6, 12, and 24 months. Abstract medical records in hospitals and doctors' offices for diagnosis, treatment, and followup/recurrence of patients. Plan for a 5-year followup to measure recurrence and management of clinical sequelae. Included insurance, income, bladder, bowel, and sexual functioning prior to diagnosis, at 6, 12, and 24 months. Sample includes 3,500 men.
SEER-Medicare Database — NCI	1986-1994	The SEER-Medicare linked data include persons reported to the SEER registries who also have been identified from Medicare's master enrollment file as being Medicare beneficiaries. The SEER-Medicare data include SEER information as well as Medicare demographic and eligibility information for all persons. For Medicare beneficiaries with fee-for-services, the SEER-Medicare data include utilization of Medicare covered health services (data about use of specific health services for Medicare beneficiaries enrolled in HMOs are not available). This utilization includes all hospital inpatient and outpatient care, physicians services, home health, hospice, and skilled nursing facility use for the entire time that the person was eligible for Medicare. For persons appearing in the SEER data who were age 65 and older, 93 percent were matched in Medicare data. The SEER-Medicare data include information for approximately 911,000 persons diagnosed with cancer between 1973 and 1993. Linkage of cases diagnosed through 1996 and Medicare claims through 1998 is in progress.  The SEER-Medicare data offer the opportunity to track SEER cases longitudinally. Thus, the data can be used to assess screening and detection performed pre-diagnosis as well as post-diagnostic surveillance. The codes used in the data are not always precise and it is difficult, at times, to distinguish screening versus diagnostic procedures.
<b>Diagnosis</b>		
SEER-Medicare Database — NCI	1986-1994	See Screening/Detection  The SEER-Medicare data can be used to assess initial treatment. SEER-Medicare augments the SEER data in that the SEER data captures only the most invasive surgical treatment while the Medicare data captures all surgical treatments in the peri-diagnosis period.
Prostate Cancer Outcomes Study — NCI	Diagnosed 10/94 through 10/95	See Screening/Detection

Survey/Database and Sponsoring Organization	Years Covered	Brief Description
<b>Treatment</b>		
Patterns of Care Study — NCI	Diagnosed 1987-1989	NCI has a Congressional mandate to report biannually on the dissemination of state-of-the-art therapy into community practice. Physician verified treatment for a sample of early stage breast (sample of about 1,000 cases each year) and B2 and C colorectal cancer (about 500 cases each year). These were used as controls for a CCOP study. Also included insurance status and protocol status.
Patterns of Care Study — NCI	Diagnosed 1990	Physician verified treatment of early stage breast (sample of nearly 1,900 cases) and B2 and C colorectal cancer (nearly 1,200 cases). Included comorbidities, insurance, hospital characteristics and protocol status.
Patterns of Care Study — NCI	Diagnosed 1991	Physician verified treatment of early stage breast (sample of nearly 1,900 cases) and B2 and C colorectal (about 700 cases) and ovarian cancer (nearly 800 cases). Included comorbidities, insurance, hospital characteristics, protocol status, and additional items on biomarkers.
Patterns of Care Study — NCI	Diagnosed 1995	Physician verified treatment of in situ and early stage breast cancer (2,500 cases), B2 and C colorectal cancer (1,000 cases), bladder cancer (670 cases), and melanoma (330 cases). Over sampled blacks and Hispanics. Included comorbidities, insurance, hospital characteristics, protocol status, biomarkers, and bone marrow transplants.
Patterns of Care Study — NCI	Diagnosed 1996	Physician verified treatment of non-small cell lung cancer (nearly 1,100 cases), ovarian cancer (over 800 cases) and melanoma (765 cases) in adults. Hodgkin's, acute lymphoblastic leukemia, acute myeloid leukemia, osteosarcoma, and Ewing's sarcoma in children ages 0-19 (total of 533 cases). Over sampled blacks and Hispanics. Included comorbidities, insurance, hospital characteristics, protocol status, biomarkers and for children the medical subspecialty.
Prostate Cancer Outcomes Study — NCI	Diagnosed 10/94 through 10/95	Study outcomes in patients newly diagnosed with prostate cancer. Focus on practice patterns and quality of life. Patient interviews at 6, 12, and 24 months. Abstract medical records in hospitals and doctors' offices for diagnosis, treatment, and followup/recurrence of patients. Plan for a 5-year followup to measure recurrence and management of clinical sequelae. Included insurance, income, bladder, bowel, and sexual functioning prior to diagnosis at 6, 12, and 24 months.
<b>Survivorship/Secondary Prevention</b>		
SEER-Medicare Database — NCI	1986-1994	See Screening/Detection Because the SEER-Medicare database contains longitudinal data for persons with cancer, the data can be used to assess long-term sequelae from treatment. The data only include information for persons age 65 and over, so they cannot be used to monitor long-term sequelae from pediatric cancers or cancers occurring in young adults.
Health, Eating, Activity, Lifestyle, and Breast Cancer Prognosis (HEAL) — NCI	1996-1999	Investigate potentially modifiable prognostic factors for stage 1-IIIa breast cancer and identify the prevalence of these risk factors in a sample of early stage breast cancer cases. Data on anthropometry, body composition, fat distribution, physical activity, and diet as well as known breast cancer risk factors, a limited hormone profile, and selected tumor related prognostic factors are being collected on a sample of women with newly diagnosed (within 6-8 months) stage 0-IIIa breast cancer. Abstract medical records in hospitals and doctors' offices for diagnosis, treatment, and followup/recurrence of patients. Plan for a 2 and 5 year followup. The current expected sample size by the end of 1999 is 850, about a quarter Hispanic, the remainder non-Hispanic whites.

Survey/Database and Sponsoring Organization	Years Covered	Brief Description
<b>Second Occurrence</b>		
SEER-Medicare Database — NCI	1986-1994	See Screening/Detection The SEER-Medicare data can be used to assess treatment for a second occurrence. Moreover, the SEER-Medicare data have the potential for identifying some of the persons with recurrences or metastasis. This information can be imputed from surgeries, radiation therapy, or chemotherapy beyond the peri-diagnostic period.
<b>Terminal Care/Death</b>		
SEER-Medicare Database — NCI	1986-1994	See Screening/Detection The SEER-Medicare data contain information on all care provided until the end of life, including specific information about hospice care.



# D

## Appendix D: List of Acronyms

ACoS	American College of Surgeons	CSRP	Cancer Surveillance Research Program
ACR	American College of Radiology	CTR	Certified Tumor Registrars
ACS	American Cancer Society	DBIR	Data-Based Intervention Research
ARB	Applied Research Branch	GIS	Geographic Information System
ASSIST	American Stop Smoking Intervention Study for Cancer Prevention	HCFA	Health Care Financing Administration
ATSDR	Agency for Toxic Substances and Disease Registry	HEAL	Health, Eating, Activity, and Breast Cancer
BCSC	Breast Cancer Surveillance Consortium	HMO	Health Maintenance Organization
BRFSS	Behavioral Risk Factor Surveillance System	HNPCC	Heritable Nonpolyposis Colon Cancer
BSA	Board of Scientific Advisors	IARC	International Agency for Research on Cancer
BSE	Breast Self Examination	MQSA	Mammography Quality Standards Act
CATI	Computer-Assisted Telephone Interview	NAACCR	North American Association of Central Cancer Registries
CCPRG	Cancer Control Program Review Group	NCCCS	National Coordinating Council for Cancer Surveillance
CCS	Cancer Control Supplement	NCCDPHP	National Center for Chronic Disease Prevention and Health Promotion
CDC	Centers for Disease Control and Prevention	NCHS	National Center for Health Statistics
CINA	Cancer Incidence in North America	NCI	National Cancer Institute
CISNET	Cancer Intervention and Surveillance Modeling NETWORK	NCRA	National Cancer Registrars Association
COC	Commission on Cancer	NFCS	Nationwide Food Consumption Survey
CPS	Current Population Survey	NHANES	National Health and Nutrition Examination Survey
CRN	Cancer Research Network	NHIS	National Health Interview Survey
CSFII	Continuing Survey of Food Intakes by Individuals	NPCR	National Program of Cancer Registries
CSB	Cancer Statistics Branch	NSMF	National Survey of Mammography Facilities

PA	Program Announcement
PCE	Patient Care Evaluation
PCOS	Prostate Cancer Outcomes Study
RFA	Request for Applications
SCLD	State Cancer Legislative Database
SEER	Surveillance, Epidemiology, and End Results
SES	Socioeconomic Status
SIG	Surveillance Implementation Group
TUS	Tobacco Use Supplement
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey
YRBS	Youth Risk Behavior Survey

