National Program of Cancer Registries
Cancer Surveillance System

Application for OMB Clearance

November 12, 2006

Point of Contact: Hannah K. Weir, Ph.D.
Centers for Disease Control and Prevention
National Center for Chronic Disease Prevention and Health Promotion
Division of Cancer Prevention and Control
1600 Clifton Road NE, Mailstop K-53
Atlanta, GA 30333
(770) 488-4783
FAX (770) 488-4759

Table of Contents

National Program of Cancer Registries
Cancer Surveillance System

A. JUSTIFICATION

A1. Circumstances Making the Collection of Information Necessary

The Centers for Disease Control and Prevention (CDC) is requesting OMB approval for the extension of the 'National Program of Cancer Registries – Cancer Surveillance System (NPCR-CSS) (OMB Control No. 0920-0469). Since 2000, state and territorial cancer registries have been collecting and reporting cancer incidence data to CDC.

Cancer is a substantial public health burden. In 2005, the American Cancer Society (ACS) estimated that more than 1.4 million Americans were diagnosed with cancer and more than 570,000 died of the disease, more than 1,500 each day (1). A disproportionate number of these deaths occur in minority and low-income groups. A substantial number of these cancers and resulting deaths could be avoided through primary prevention (e.g., smoking cessation, diet, exercise) or through early detection and treatment (e.g., mammography, Pap smears).

It is estimated that 9.4 million Americans are currently alive with a history of cancer. The National Institutes of Health estimates the cost of cancer is about $210 billion including ($74 billion) direct costs to treat cancer and ($136 billion) indirect costs in lost productivity due to illness and premature death (2).

The Centers for Disease Control and Prevention (CDC) and the states face the challenge of reducing cancer morbidity and mortality through prevention and early detection. Within CDC, the Division of Cancer Prevention and Control (DCPC) plans, directs, and supports cancer control efforts through collaboration with prevention partners in state health agencies; federal agencies; academic institutions; and national, voluntary and private sector organizations. To obtain a firm basis for such programs, DCPC is actively involved in surveillance and applied research.

Recognizing the public health value of comprehensive cancer surveillance at the state and national level, Congress mandated the National Program of Cancer Registries (NPCR) in 1992 by enacting the Cancer Registries Amendment Act, Public Law 102-515 **(Attachment 1**); the law was reauthorized in 1998. This legislation authorizes the CDC to provide funds to states and territories to: 1) improve existing cancer registries; 2) plan and implement registries where none existed; 3) develop model legislation and regulations for states to enhance the viability of registry operations; 4) set standards for data completeness, timeliness, and quality; 5) provide training for registry personnel; and 6) help establish a computerized reporting and data-processing system.

In fiscal year 2005, CDC awarded $37 million to fund 45 states, three territories, and the District of Columbia for central cancer registries operations **(Attachment 2).** The National Cancer Institute (NCI) funds the remaining 5 states (Utah, Connecticut, Iowa, New Mexico, and Hawaii). Four states (California, Kentucky, Louisiana and New Jersey) receive funding from both CDC and NCI.

Information is collected and maintained at CDC under Section 306 of the Public Health Service (PHS) Act [42 USC 242(k)] (**Attachment 3**).

A2. Purpose and Use of the Information Collection

The National Program of Cancer Registries-Cancer Surveillance System (CSS) is designed to provide cancer incidence data that meet CDC's responsibilities for public health surveillance while enhancing the quality, completeness, and timeliness of state cancer incidence data and monitoring progress toward the NPCR program objectives.

In the first five years of the program, the grantees were funded to improve the completeness, timeliness, and quality of population-based central cancer registry data. Under the current National Cancer Prevention and Control program announcement (CDC program announcement #02060), grantees are requested to submit annual cancer incidence data to CDC.

As stated in Public Law 102-515, state central cancer registries must collect each form of invasive cancer (with the exception of basal cell and squamous cell carcinoma of the skin). The central cancer registry routinely receives a standard set of data items on all cancer patients diagnosed in the state from hospitals, pathology labs, clinics and private physicians. Based on negotiations with the state, larger facilities may report monthly to the central cancer registry and smaller facilities less frequently. NPCR has established a goal of no more than six months between the diagnosis of cancer and receipt by the central registry. The cancer registries maintain these data items permanently in longitudinal databases that are used for public health surveillance, program planning and evaluation, and research.

Once a year in January, in lieu of a quarterly report, CDC requests cumulative data from central cancer registries beginning with their reference year for NPCR (1995 for most programs) to one year after the close of the most current diagnosis year (e.g., diagnosis 1995-2004 data in the calendar year 2006). **Attachment 4** is a copy of the submission specifications that went to NPCR grantees in October 2005 providing instructions for the reporting of cancer incidence data to CDC in January 2006. Attachment 2 of the document contains a list of data items that NPCR grantees are required to report**.** CDC updates its longitudinal database each year with data from the most recent diagnosis year from the states**.**

A data contractor, ORC Macro (Macro), has been retained to assist with data management and analysis. Based on annual CSS submissions, standardized reports are generated by Macro for the grantees and the CDC. These reports allow the program to monitor and evaluate the grantees performance with respect to the quality and completeness of their data. Data will be used by CDC for program planning and improvement and CDC will provide regular feedback to grantees based on their data submission and will tailor technical assistance as indicated. In particular, CDC monitors the ability of each grantee to reach data standards with respect to the completeness, timeliness and quality of the data.

- Within 24 months past the close of the most recent diagnosis year, each NPCR grantees is expected to have registry data that include at least 95% of the expected, unduplicated cases where the expected cases are estimated by using methods developed by NAACCR (3). Because some cancer patients receive diagnostic or treatment services at more than one reporting facility, cancer registries perform a procedure known as *unduplication* to ensure that each cancer case is counted only once (4). Within 12 months past the close of the diagnosis year, grantees are expected to have registry data that include at least 90% of expected cases.
- Within 24 months past the close of the diagnosis year, no more than 3% of cases are to have been ascertained solely on the basis of a death certificate. The proportion of cases ascertained solely on the basis of a death certificate, with no other information on the case available after the registry has completed a routine procedure known as "death clearance and follow back" (5) is an approximate measure of the completeness of case ascertainment.
- Within 24 months past the close of the most recent diagnosis year, each NPCR grantee is expected to have registry data with no more than 2% of cases having missing information on sex; no more than 2% of cases having missing information on age; no more than 3% of cases having missing information on race; and no more that 2% of cases having missing information on county of residence.
- Within 24 months past the close of the most recent diagnosis year, each NPCR grantee is expected to have registry data where at least 99% of the registry's records passed a set of single-field and inter-field computerized edits. Computerized edits are computer programs that test the validity and logic of data components. Within 12 months past the close of the diagnosis year, grantees are expected to have data where at least 97% of the record pass edits.

These performance indicators may be modified or changed over time to more accurately reflect program priorities and areas of concerns. These performance indicators will also be used for reporting to CDC officials, Congress and other national stakeholders.

Data from NPCR grantees who meet program standards may be used for the following:

- **Cancer Surveillance** The CDC and the states face the challenge of reducing cancer morbidity and mortality through prevention and early detection. Effective control of chronic diseases, such as cancer, requires the regular, ongoing collection and analysis of health-related data to monitor the frequency and distribution of disease in the population. The NPCR-CSS will help CDC continue to meet its public health responsibilities by providing routine surveillance reports on the national cancer burden by demographic characteristics, tumor characteristics, survival time, and other items of interest to the public health agencies responsible for the design, implementation, and evaluation of cancer prevention and control activities. CDC's prevention efforts will be enhanced by the ability to target areas with high rates of cancer with appropriate screening such as mammography, Pap smears, and colorectal cancer screening. The Agency for Healthcare Research and Quality (http://www.ahrq.gov/) includes measures for effectiveness of care in cancer. The *2005 Healthcare Quality Report now* includes rates of advanced stage female breast, colorectal and cervical cancer by state.

  For the past five years, CDC and the National Cancer Institute, in collaboration with the North American Association of Central Cancer Registries (NAACCR) have published *United States Cancer Statistics (USCS) (http://www.cdc.gov/CANCER/npcr/uscs/index.htm).* The USCS report contains a set of official federal cancer incidence statistics from each state that had high quality registry data. The report provides state-specific and regional data for cancer cases diagnosed in 2002, the most recent year for which federal data is available (**Attachment 5)**. This past year, forty-five statewide population-based cancer registries and the District of Columbia, covering 93% of the U.S. population, met the eligibility criteria for inclusion in this report. Data for selected cancer sites is also available as pre-calculated counts and rates on the NCI/CDC State Profiles Website (http://statecancerprofiles.cancer.gov/) and on the CDC's Wonder Website (http://wonder.cdc.gov/CancerIncidence.html).

  The Council of State and Territorial Epidemiologists Association (CSTE) has voted to include cancer as part of the chronic disease indicators of the National Public Health Surveillance System (NPHSS) (6). The NPCR-CSS continues to work with other surveillance systems such as HIV/AIDS, Birth Defects Monitoring Program, Pregnancy Risk Assessment Monitoring System and the Childhood Blood Lead Surveillance Program to make timely data available for the NPHSS and publication in the Morbidity and Morality Weekly Report.

- **Program Planning and Evaluation** CDC sponsors and supports a wide variety of public health programs in the U.S. designed to monitor and reduce morbidity and mortality from cancer such as the National Comprehensive Cancer Control Program, National Tobacco Control Program, the National Breast and Cervical Cancer Early Detection Program, the National Colorectal Cancer Roundtable, prostate cancer

control initiatives, and the National Skin Cancer Prevention Education Program. Increasingly, there is Congressional and public demand for federal agency documentation and accountability of achievement of program objectives and outcomes (e.g., the Government Performance and Results Act of 1993).

Cancer information collected under NPCR will be very important to evaluate the success and remaining challenges in meeting CDC program goals and objectives, as well as to identify areas that could benefit from education and training, technical assistance, and other resources.

- **Research** When all NPCR-funded cancer registries meet the data criteria for publication in United States Cancer Statistics (currently 40), the registries will provide geographic coverage for 96% of the U.S. population **(Attachment 2).** (The remaining 4% of geographic coverage will come from the NCI-funded cancer registries.) State registries, with the exception of large densely populated states, lack the number of cases to permit calculation of stable rates for special populations and in some cases the general population. Currently available data are frequently inadequate for the surveillance of cancer in special populations such as racial and ethnic minorities, medically under-served groups, and populations at high risk for selected cancers that may not be identifiable in statewide databases because of small numbers or other special circumstances.

  Public use and restricted access datasets are in development (**Attachment 6**) that will provide a statistical basis for analyzing the cancer burden on a regional and national level (http://www.cdc.gov/cancer/npcr/datarelease.htm).

A3. Use of Improved Information Technology and Burden Reduction

The NPCR requires states to report only those data items specified in the law and needed for research and program planning and evaluation **(Attachment 4**). After data collection and editing have been completed, the state cancer registry data are sent to the CDC electronically using standard data definitions and record layouts. The current data definitions and record layout **(Attachment 7)** are recommended by the North American Association of Central Cancer Registries (NAACCR) which includes representatives from grantees and the CDC, as well as other cancer reporting organizations (7)**.** Most grantees have been using these standards since they were established for inter-state data exchange and for reporting to NAACCR. The use of existing standards helps reduce errors and the electronic transmission of data will be efficient and minimize the reporting burden on the states.

A4. Efforts to Identify Duplication and Use of Similar Information

At the national level, cancer incidence data are available through the National Cancer Institute's (NCI) SEER Program, which represents 9-26% of the population of the U.S. The SEER database contains information on 9% of the population on cancers diagnosed between 1973 and 2002 (or the most recent year available) from nine areas including the states of Connecticut, Iowa, New Mexico, Utah, Hawaii and the metropolitan areas of Detroit, San Francisco, Seattle-Puget Sound, and Atlanta (http://seer.cancer.gov/). In 1992, Los Angeles and San Jose/Monterey were added bring the coverage to 11%. And in 2002, four additional states were added to the SEER Program (California, Kentucky, Louisiana, New Jersey) bringing SEER total coverage to 26% of the U.S. population. These four states receive joint funding from the two federal programs and report their data to both federal agencies. SEER data are of high quality and are used to analyze long term trends in cancer incidence, patient survival, and for many other research purposes. While the SEER data are appropriate for

analyses of major cancers in large population subgroups, they are not always adequate for analysis of U.S. regions, minority populations and rare cancer analyses. These data are not useful for most states for program planning and evaluation. When all NPCR-funded registries meet NPCR data standards, the NPCR Cancer Surveillance System will cover 96% of the United States and will complement the SEER data to provide 100% coverage of the US population. In the three states where the SEER program covers a part of the state (Georgia, Michigan, Washington) and the state participates in the NPCR, there is no duplication of effort. The SEER program reports data from their catchment area to the NPCR-funded state central cancer registry.

NAACCR plays a leadership role in setting standards for the collection of cancer data and currently publishes population-based state cancer incidence data and aggregated state data yearly in Cancer Incidence in North America (CINA) (8). The submission of data to NAACCR is voluntary and varies from year to year. No public use data set is available to meet both public health surveillance needs and NPCR needs for program planning and evaluation.

The National Cancer Data Base (NCDB) from the American College of Surgeons (ACoS) contains data items required by the Commission on Cancer Approvals Program. NCDB is based on approximately 1,300 participating hospitals. The program is in its seventeenth year of national operation and approximately 70% of all U.S. cancer cases (850,000) are collected annually. The data are not population-based since NCDB does not collect all cancer cases in a defined geographic area and cannot be used to calculate incidence rates.

While there are a number of cancer registration activities in the U.S., it is clear that the resulting data do not meet the public health need for a national cancer surveillance system. The NPCR-CSS is unique in meeting the national need for a population-based dataset with adequate numbers of rare cancers, representation of minority populations, and state-based data for program planning and evaluation.

A5. Impact on Small Businesses or Other Small Entities

No small businesses will be involved in this study.

A6. Consequences of Collecting the Information Less Frequently

The NPCR-CSS data aggregation occurs on an annual basis in January in place of a quarterly written report. The ability of CDC to monitor and improve program effectiveness would be compromised if data were collected less frequently. It is essential that CDC and State program managers evaluate program strengths and weaknesses on an annual basis and make adjustments. It is also important to provide annual information on the national cancer burden to CDC officials, Congress, constituents, and other Federal, State, and local agencies.

There are no legal obstacles to reduce the burden.

A7. Special Circumstances Relating to the Guidelines of 5 CFR 1320.5

There are no special circumstances contained within this application.

A8. Comments in Response to the Federal Registrar Notice and Efforts to Consult Outside the Agency

A. A 60-Day Federal Register Notice was published on May 9, 2006 in Volume 71, No. 89 pp. 26969-70 **(Attachment 8).** One public response (see Attachment 8 for response) was received. No changes were made to the proposed project based on this response, as the public comment did not relate to the utility and scope as proposed.

B**. Attachment 9** contains a list of experts in cancer registration that met with NPCR staff on August 12, 1998 to provide expert advice on data aggregation. These experts include representatives from grantees, NAACCR, NCI, and the American Cancer Society. There were no major problems to be resolved.

In May 1999, NPCR distributed a Rationale and Approach Paper for NPCR-CSS to states and national partners (e.g., ACS, NCI, NAACCR) and comments were solicited. The most frequently asked question was about confidentiality of data. Some states have legislation that restricts the exchange of data and some states have policies that discourage the practice. The CDC respects state laws governing data release, and will work with states on this issue. In response to these concerns, CDC applied for and received a Confidentiality Assurance**.** CDC has based its approach to confidentiality for NPCR-CSS on that of the National Center for Health Statistics (NCHS). The NCHS has been successful in protecting confidential health data for more than ten years. An NCHS Confidentiality Expert has reviewed our data release policy.

A9. Explanation of Any Payment or Gifts to Respondents

No payment will be made to respondents (grantees) to submit NPCR data to CDC.

A10. Assurance of Confidentiality Provided to Respondents

The CDC Privacy Act and Confidentiality Officer has reviewed this application and has determined that the Privacy Act is not applicable. Although grantees have access to personally identifiable information, only de-identified records are transmitted to CDC. Additional information on privacy safeguards applicable to data collection, de-identification, coding, transmission, storage, and reporting appears below.

Confidentiality and privacy are of paramount concern to the NPCR because of the confidentiality concerns of the grantees, the private nature of medical data in a cancer surveillance database, and the potential for direct and deductive identification of an individual in the NPCR-CSS. After extensive discussions with the CDC Privacy Officer, CDC obtained an Assurance of Confidentiality (308(d))

on June 7, 2000 **(Attachment 10**). The proposed new data collection will be covered under an extension of the 308(d).

The threat of direct identification of an individual in NPCR-CSS data is remote because personal identifying data (name, social security number and street address) will not be reported to the CDC. However, a unique identifier assigned by the state to each individual cancer patient will be reported to CDC. While each record constitutes a single primary cancer, it is necessary to identify multiple primary cancers in an individual. The grantees maintain the linkage information between the unique codes and the personal identifies in their database in order to respond and follow-up on data queries from CDC. Since multiple primary cancers are a matter of research interest, the public use files must also contain a unique identifier.

Of greater concern is the geographic data (e.g., county, census tract, zip code) that will be reported to CDC and the potential for deductive identification. Geographic data could be combined with other publicly available information and potentially be a threat to confidentiality. Because surveillance and analysis of cancer by county are of public health interest, NPCR proposes to make these data available, but to limit access, require a signed data release agreement, and provide guidelines for data use. CDC will create multiple datasets of increasing sensitivity with respect to geographic data **(Attachment 6).** In the first tier of data (the least confidential), state would be the smallest geographic unit released. More sensitive dataset would contain county level data. The user would have to describe the need for county level data. In data tiers one and two, other potential identifiers include date and place of birth, race, vital status, date of last contact and rare primary sites. These data will be examined prior to release, and if necessary recoded to protect small population subsets. For example, only the year will be provided for potentially identifying dates such as date of birth, diagnosis and death. Once tier one and two data have been examined and recoded, we believe that they will not pose a significant risk to confidentiality.

A third and most sensitive dataset would contain census tract and zip code in addition to the variables in the first and second datasets. This dataset is the most likely to create opportunities for deductive identification and as such, CDC intends to guard this dataset very carefully. To provide data, CDC would need a research protocol, local IRB approval, and a plan to assure confidentiality. If CDC staffs were co-investigators with states on the analysis, an IRB protocol would be submitted to a CDC IRB for review. Data would be provided to meet specific needs and data items would be collapsed when necessary to protect confidentiality. Only a limited number of tier three analyses would be approved each year. The NPCR-CSS data use agreement is based on the NCHS model **(Appendix 6, Attachment 10)**.

To address the issue of deductive identification of an individual because of small numbers (e.g., in a census tract), guidelines from the NCHS Staff Manual on Confidentiality will be used (9). NCHS has guidelines for published data and one for micro data files or public use files. The guidelines for published data include: 1) “In no table should all cases of any line or column be found in a single cell”, 2)

"In no case should the total figure for a line or column of a cross-tabulation be less that 3", and 3) "In no case should a quantity figure be based upon fewer than three cases." The guidance for avoiding inadvertent disclosures through the release of micro data tapes includes 1) "The tape must not contain any detailed information about the subject that could facilitate identification and that is not essential for research purposes (e.g., exact date of the subject's birth)" and 2) "Geographic places that have fewer than 100,000 people are not to be identified on the tape." These guidelines from NCHS will serve as a model for CDC as confidentiality procedures are established. In addition, the program will need to be attentive to changes in the environment that may impact efforts to maintain confidentiality.

The CSS data are sent to a contractor (ORC Macro). CDC staff and the contract staff have developed a security plan. This security plan ensures that the data are kept secure and confidential.

- The NPCR-CSS project data reside on a dedicated server at ORC Macro. To ensure the security and confidentiality of project data, the following provisions have been incorporated in the ORC Macro NPCR-CSS Security Plan.
- The NPCR-CSS server is housed in a secure facility at ORC Macro's Bethesda office with a guard on duty in the lobby 24 hours a day. Elevator and stairwell access is controlled by card key. The server resides on ORC Macro's local area network (LAN) behind ORC Macro's firewall.
- Access to the NPCR-CSS server is limited to authorized ORC Macro project staff (see attachment). It is password protected on its own security domain. No one, including non-project staff at ORC Macro, is allowed access to the NPCR-CSS.
- All ORC Macro project staff must sign a confidentiality agreement before passwords and keys are assigned. All staff must pass background checks appropriate to their responsibilities for a public trust position.
- NPCR-CSS data that are submitted electronically are encrypted during transmission from the states. They arrive on a document server behind ORC Macro's firewall. Each State has its own directory location so no State has access to another State's data. The data are moved automatically from the document server to the NPCR-CSS server.
- Receipt and processing logs are maintained to document data receipt, file processing, and report production. All reports and electronic storage media containing NPCR-CSS data will be stored under lock and key when not in use and will be destroyed when no longer needed.
- A comprehensive security plan has been developed by ORC Macro's security team. The security team consists of: June Bray-Business Steward, Kevin Zhang-Technical Steward, Leo Shen-Security Officer, David Radune-Database Administrator, and Gretchen Sinclair-WAN & LAN Security Steward. All project staff receive annual security awareness training covering security procedures. The ORC Macro project security team oversees operations to prevent unauthorized disclosure of the NPCR-CSS data.
- Periodic review and update of ORC Macro security processes is conducted to adjust for rapid changes in computer technology and to incorporate advances in

security approaches. The security plan will be amended as needed to maintain the continued security and confidentiality of NPCR-CSS data.

- At CDC, all CSS datasets are maintained for restricted access on a secure LAN server. Access to these datasets is only granted when appropriate confidentially release forms have been signed and returned to the CSS Data Security Steward: Joseph D. Rogers.

The study protocol (#2594) for CSS has been reviewed and approved by a CDC Institutional Review Board (IRB). The most current notice of approval (October 12, 2006) is attached **(Attachment 11).**

A11. Justification for Sensitive Questions

This data collection includes sensitive information about cancer diagnosis and treatment, which is central to the purposes of the project. In addition, Race and Ethnicity data are collected per HHS guidelines, and for use in epidemiologic analyses.

A12. Estimation of Annualized Burden Hours and Costs

A. These data are already collected and aggregated at the state level. Thus the additional burden on the states should be small and should only involve the time to electronically submit the data. Program implementation would require funded states to report data to the CDC on an annual basis twelve months after the close of a diagnosis year and again at twenty-four months to obtain more complete incidence data and vital status from mortality data. The burden of reporting data to CDC is reduced by the use of data standards adopted by all NAACCR member registries as detailed in section A3 of the Supporting Statement.

Table A.12-A Number of Respondents and Estimated Burden Hours.

Respondents	Number of Respondents	Frequency of Responses	Average Burden per Response (in hours)	Total burden in hours
State Departments of Health	63	1	2	126

States prepare their data files and send them electronically to CDC. The web page displays the OMB control number, the expiration date and a burden statement **(Attachment 12).** This information appears on the log in page of the website that the states use to transfer their files electronically

B. The annualized cost to respondents for the 2.0 hour burden of reporting data to CDC is estimated to be $3,465 per year among the 50 states, 12

US territories and the District of Columbia. It is estimated that the following state cancer registry personnel will be required to help prepare and submit data electronically to CDC: data managers, and information technology and clerical staff. The hourly wage rates are averages of the wages paid to state health department personnel participating in the NPCR.

Table A.12-B. Annualized cost to respondents.

Respondents	Number of Respondents	Hours to Respond	Hourly Wage Rate	Respondent cost)
Data Managers	63	1.0	$30	$1,890
Information technology staff	63	0.5	$35	$1,102
Clerical staff	63	0.5	$15	$473
Total				**$3,465**

A13. Estimates of Other Total Annual Cost Burden to Respondents and Record Keepers

The computer hardware and software needed for an electronic data submission to CDC are readily available to grantees since they collect and distribute cancer incidence data for state purposes; hence no capital or maintenance costs are anticipated.

A14. Annualized Cost to the Federal Government

The average annual cost for the contractor for data collection is $1,101,215 per year for a five-year total of $5,506,078. A data management contract was awarded to ORC Macro in 2000. Additional annual costs include personnel costs of federal employees involved in oversight and analysis. The annual staff cost is estimated at $120,000 (1 epidemiologists FTE, 0.2 public health advisor FTE, and miscellaneous expenses include travel, copying, etc.).

Table A14. Estimated Annualized Federal Government Cost Distribution:

A15. Explanation for Program Changes or Adjustments

No changes in hour burden.

A16. Plans for Tabulations and Publication and Project Time Schedule

CDC requests a 3-year clearance for the proposed, recurring data collection. Data will be received every year in January from grantees. In addition to data from the current diagnostic year, data will be requested back to the reference year for the program, which for most states is 1995. Data submissions will usually be a combination of new data from the most recent diagnostic year and re-submissions from previous years that are improved in quality and completeness. Each year the process of data submission, data editing, data enhancement, and creation of public use datasets will be repeated (Table A16). The schedule each year will be:

Table A16. Time Schedule for Data Reporting, Analysis and Publication:

A17. Reason(s) Display of OMB Expiration Date is Inappropriate

There is no request for a date display exemption.

A18. Exceptions to Certification for Paperwork Reduction Act Submission

There are no exceptions to the certification.

B. STATISTICAL METHODS

B1. Respondent Universe and Sampling Methods

Respondents are the 49 states and territories that currently receive CDC funds from the NPCR to create and enhance central cancer registries. The maximum number of respondents possible is 63, which include all 50 states, 12 territories, and the District of Columbia that are eligible to apply for funds from NPCR.

Statistical methods are not employed. Data collection at the state level is population based and these data will be reported annually to NPCR-CSS. The number of years of data that a state will report depends on when funding began (1995 at the earliest), whether a state had an established registry (an enhancement state) or had to begin a new registry (a planning state), and when the state began a population-based central cancer registry (the reference year). Twenty-seven NPCR states have a reference year of 1995, when they began collecting cancer incidence data as part of the NPCR, fourteen have a reference year of 1996, and three a reference year of 1997. Two additional states and the three U.S. territories have reference years between 1998 and the year 2000. **Attachment 13** lists the number of cases each state reported to NPCR-CSS in 2005. In diagnosis year 2002, over 1.4 million incident cases of cancer were reported to CDC from 45 states and the District of Columbia.

B2. Procedures for the Collection of Information

Required program data will be reported to CDC by grantees once a year in January in lieu of one quarterly report. Since the grantees collect and aggregate data for local public health purposes, they have the primary responsibility for information collection procedures. As depicted in **Attachment 14**, Data Collection and Processing Flow Chart, the first step occurs when a physician makes a diagnosis of cancer. Once a definitive diagnosis has been established and treatment planned, the data are entered into a computer, usually with a commercial software package that includes quality control measures to assure high quality data (step 2). NPCR has established a goal of no more than six months between the diagnosis and computerization of cancer data. Step three on the flow chart occurs when hospitals and clinics perform additional quality control measures over and above what is performed at data entry and send data to the central cancer registry (step 4). Quarterly submissions to the

central registry are common, but larger facilities may report more often, and smaller facilities, less frequently.

After the central cancer registry receives the provider data, each incoming case must be checked against the existing database to ascertain if it is a new case or has been reported previously. At the same time additional quality control measures are applied (step 5). Based on this processing, the central registry may return data to the reporting facility for clarification (step 6). Once quality control standards are met and the data are complete, they are ready for use and dissemination by the state and submission to CDC (steps 7 and 8). This process usually takes 12 to 18 months (12 months is the goal) after the close of the year in which the cancer is diagnosed.

Once the CDC receives the data, they are processed and data evaluation reports are generated as indicated in step nine on the flow chart. The data evaluation reports **(Attachments 15)** include the results of evaluating state data by the data standards for completeness of case ascertainment and data quality as adopted by NPCR for program goals and publication (Standard Status Report) and a report detailing the states' submission (Submission Summary Report) including details of edit errors.

When standards of completeness and quality have been met, CDC will aggregate state data and make them available in non-confidential recalculated rates on the Internet in a format that facilitates obtaining data by sex, race, age, and other common factors of interest**.** Public use datasets will be made available to researchers with a signed data release agreement. Any data published from NPCR-CSS, whether in surveillance reports, on the Internet, or in public use datasets will be scrutinized to assure that the confidentiality of the individual is protected.

In addition to the activities described above, other NPCR program activities impact completeness of data for NPCR-CSS. Grantees are funded for inter-state data exchange to obtain cancer data on residents who travel to other states for diagnosis or treatment. Grantees must also link state incidence data with state mortality data to obtain cases that are first diagnosed at death (death certificate only cases). In 2001, 39 states had a law and/or regulations in place that meet all eight criteria as specified in Public Law 102-515 (Hutton, 2001). As of 2004, 45 states have such a law. Prior to the advent of the NPCR, only eight states had authorizing legislation in place that mandated the reporting of cancer to the central cancer registry from all facilities and practitioners that provide screening, diagnostic and treatment services. In some states, this made complete case ascertainment very difficult.

A major program activity that has impacted data quality has been the development of EDITS, a PC-based software product for editing data. CDC collaborated with NCI, ACoS, ACS, and NAACCR to create edit routines for standard cancer data items. The NPCR has established a list of data edits for NPCR grantees to run on their data using the EDITS software. These edits verify each data item reported to CDC as well as important interfield and inter-record relationships (10). The same edits that grantees run will be run on data reported to NPCR-CSS so that data quality can be assessed uniformly for each program participant.

B3. Methods to Maximize Response Rates and Deal with Non response

In January 2006, 48 of 49 eligible NPCR grantees (45 states, the District of Columbia, and 2 U.S. territories) reported their data to NPCR-CSS. The use of existing data standards and record layouts for electronic submission of data makes it easy for states to comply with the request. Many NPCR states submit data to NAACCR for CINA and exchange data with neighboring states using these standards and formats. There should be few technical difficulties for states in using these familiar processes.

In addition, to ease reporting, there are a number of other incentives for states to submit data. The incentives include an independent and detailed assessment of data quality and the recoding of important data items such as primary site and histology to national standards used for analysis

There is no reason to believe that the response rate in subsequent years would be much lower than 100%. If a state has difficulty submitting data, the CDC Project Officer and/or the CDC data contractor would provide assistance. NPCR will also be working with states to assure that they have complete coverage of the population in their catchment area.

B4. Test of Procedures or Methods to be Undertaken

Following the awarding of the data management contract in 2000, CDC requested grantees to voluntarily provide the contractor with test data that summer. In total, 31 state and territorial cancer registries send files of cancer incidence data. This enabled the contractor to create and test the programs for receiving and evaluating the completeness and timeliness of the cancer incidence data, including processes used to edit data, create reports, provide feedback, display data on the Internet, and create datasets. The first NPCR request for data was held in 2001. Each year the system is tested and refined based on test data from previous years' submissions. States are not requested to send additional data to test the system.

Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The data collection was designed by Hannah Weir, PhD, (770-488-3006) technical monitor and epidemiologist from the Cancer Surveillance Branch, Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Control, CDC. The CDC project officer for the contract is Christine Dauer, public health advisor at the same address. NPCR data collection and data quality standards are formulated and recommended by NAACCR. Staffs at NCHS were consulted extensively about their approach to Confidentiality.

List of Attachments

Attachment 1	Public Law 102-515 Cancer Registries Amendment Act
Attachment 2	Map of National Program of Cancer Registries grantees
Attachment 3	Section 306 of the Public Health Service Act [42 U.S.C. 242(k)]
Attachment 4	NPCR-CSS Submission Specification document
Attachment 5	*United States Cancer Statistics: 2002 Incidence and Mortality*
Attachment 6	NPCR-CSS Data Release Policy
Attachment 7	Record Layout and Data Dictionary (excerpted from NAACCR Data Standards and Data Dictionary, version 11)
Attachment 8	Federal Register Notice and Response
Attachment 9	Participants in Consultation Outside the Agency
Attachment 10	308(d) Assurance of Confidentiality Certificate Approval
Attachment 11	Institutional Review Board - Approval Notification
Attachment 12	Web pages containing OMB burden statement
Attachment 13	Cases Reported to CDC in Year 2005 Submission
Attachment 14	Flow Chart of Data Reporting for NPCR-CSS
Attachment 15	NPCR-CSS Data Evaluation Reports

References

1 American Cancer Society. *Cancer Facts and Figures 2005*. Atlanta (GA): American Cancer Society; 2005.

2 National Heart, Lung, and Blood Institute (NHLBI). *Fact Book, Fiscal Year 2004*. Bethesda (MD): NHLBI; 2005. Available at http://www.nhlbi.nih.gov/about/factpdf.htm

3 Tucker TC, Howe HL, Weir HK. Certification of population-based cancer registries. *Journal of Registry Management* 1999; 26(1): 24-27.

4 North American Association of Central Cancer Registries (NAACCR). *Standards for Cancer Registries:* vol. III: *Standards for Completeness, Quality, Analysis and Management of Data*. Springfield (IL): NAACCR; 2000.

5 Menck HR, West DW. Central cancer registries. In: Hutchison CL, Roffers SD, Fritz AG. *Cancer Registry Management: Principles and Practice*. Lenexa (KS): National Cancer Registrars Association; 1997. p. 395-422.

6 Chronic Disease Committee, Council of State and Territorial Epidemiologists. Inclusion of Cancer Incidence and Mortality Indicators in the National Public Health Surveillance System (Position Statement #CD 4). June 1998.

7 Havener, L and Hulstrom D editors. Standards for Cancer Registries, Vol II: Data Standards and Data Dictionary, 10nd ed. Springfield, Ill: North American Association of Central for Cancer Registries; 2004.

8 Hotes JL, Wu XC; Howe H, McLaughlin CC et al. editors. *Cancer in North America, 1998-2002:* vol. 1: *Incidence.* Springfield (IL): North American Association of Central Cancer Registries; 2005. Available at http://www.naaccr.org/index.asp?Col_SectionKey=11&Col_ContentID=50

9 NCHS Staff Manual on Confidentiality, 1997:National Center for Health Statistics, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Attachment 1

Public Law 102-515 Cancer Registries Amendment Act

Attachment 2

Map of National Program of Cancer Registries grantees

Attachment 3

Section 306 of the Public Health Service Act [42 U.S.C. 242(k)]

Attachment 4

NPCR-CSS Submission Specification document

Attachment 5

United States Cancer Statistics: 2002 Incidence and Mortality

Attachment 6

NPCR-CSS Data Release Policy

Attachment 7

Record Layout and Data Dictionary
(extracted from NAACCR Data Standards and Data Dictionary, version 11)

Attachment 8

Federal Register Notice and Response

Attachment 9

Participants in Consultation Outside the Agency

Attachment 10

308(d) Assurance of Confidentiality Certificate Approval

Attachment 11

Institutional Review Board - Approval Notification

Attachment 12

Web pages containing OMB burden statement

Attachment 13

Cases Reported to CDC in Year 2005 Submission

Attachment 14

Flow Chart of Data Reporting for NPCR-CSS

Attachment 15

NPCR-CSS Data Evaluation Reports