Information Collection Request

Reinstatement

National Program of Cancer Registries Cancer Surveillance System OMB No. 0920-0469

Supporting Statement: Part A

Program Official

Reda J. Wilson, MPH, RHIT, CTR Epidemiologist NPCR-CSS Lead Cancer Surveillance Branch Centers for Disease Control and Prevention Atlanta, Georgia 770-488-3245 (office) 770-488-4759 (fax) dfo8@cdc.gov

March 21, 2013

TABLE OF CONTENTS

ABSTRACT

A. JUSTIFICATION

- A1. Circumstances Making the Collection of Information Necessary
- A2. Purpose and Use of the Information Collection
- A3. Use of Improved Information Technology and Burden Reduction
- A4. Efforts to Identify Duplication and Use of Similar Information
- A5. Impact on Small Businesses or Other Small Entities
- A6. Consequences of Collecting the Information Less Frequently
- A7. Special Circumstances Relating to the Guidelines of 5 CFR 1320.5
- A8. Comments in Response to the Federal Registrar Notice & Efforts to Consult Outside the Agency
- A9. Explanation of Any Payment or Gifts to Respondents
- A10. Assurance of Confidentiality Provided to Respondents
- A11. Justification for Sensitive Questions
- A12. Estimation of Annualized Burden Hours and Costs
- A13. Estimates of Other Total Annual Cost Burden to Respondents or Record Keepers
- A14. Annualized Cost to the Federal Government
- A15. Explanation for Program Changes or Adjustments
- A16. Plans for Tabulations and Publication and Project Time Schedule
- A17. Reason(s) Display of OMB Expiration Date is Inappropriate
- A18. Exceptions to Certification for Paperwork Reduction Act Submissions

REFERENCES

ATTACHMENTS

- 1a Cancer Registries Amendment Act, Public Law 102-515
- 1b Section 301 of the Public Health Service Act [42 U.S.C. 242k]
- 1c American Recovery and Reinvestment Act of 2009, Public Law 111-5
- 2 Data Collection and Data Flow Process for NPCR CSS
- 3a Standard NPCR CSS Submission Specifications
- 3b Enhanced NPCR CSS Submission Specifications
- 4 NPCR CSS Data Release Policy
- 5 308(d) Assurance of Confidentiality Certificate Approval
- 6 Federal Register Notice
- 7 Participants in Consultation Outside the Agency
- 8 Institutional Review Board Approval Notification
- 9 OMB Burden Statement on NPCR CSS Web Site
- 10 Cases Reported to CDC in Year 2012 Submission
- 11 Sample of NPCR CSS Data Evaluation Report

ABSTRACT

In the United States, cancer registration is the only method for systematically collecting and reporting population based information about cancer incidence and prevalence. Cancer registration depends upon Central Cancer Registries (CCR), funded in part by CDC's National Program of Cancer Registries (NPCR). Current NPCR grantees include 45 state-based central cancer registries (CCR), the District of Columbia, Puerto Rico, and the Pacific Islands jurisdictions. In November of each year, NPCR-funded CCRs have reported specific information to CDC as a function of the National Program of Cancer Registries Cancer Surveillance System (NPCR CSS), OMB no. 0920-0469, exp. 11/30/2012. Because the CCRs need time to collect and reconcile records from a variety of sources, each submission to CDC reflects cancers diagnosed or treated during the previous one to two years.

CDC requests OMB approval to reinstate the NPCR CSS information collection, with changes. First, the frequency of reporting to CDC will be changed from an annual to a semi-annual schedule. The additional report will allow CDC to compile preliminary cancer incidence estimates and assess data quality in advance of the lengthy process of data validation required for each registry's final annual report.

The third set of changes applies to a subset of 10 central cancer registries. Through a multimillion dollar American Recovery and Reinvestment Act (ARRA) appropriation, these cancer registries received funding to develop common standards and reporting mechanisms for enhanced description of cases of breast cancer, colorectal cancer, and chronic myeloid leukemia (CML). The enhanced data items will support more in-depth analysis of treatment strategies and patient outcomes than is currently possible with the standard NPCR CSS information collection. The 10 registries that participated in the enhancement process will begin reporting the additional data items to CDC in 2013 as part of their routine submission (Appendix 3b). CDC plans to make de-identified data available for comparative effectiveness research (CER), also known as patient-centered outcomes research (PCOR).

OMB approval is requested for three years. The respondents for the NPCR CSS are the 48 CDC funded central cancer registries. The estimated burden per response is two hours because cancer incidence data are already collected, aggregated and used for analyses at the state level. This process of data collection is conducted under the authority of state laws which require reporting of cancer data to the central registry. Consequently, the additional burden of electronically reporting the information to CDC is small and the number of data items in the report does not affect the estimated burden per response.

National Program of Cancer Registries Cancer Surveillance System

A. JUSTIFICATION

A1. Circumstances Making the Collection of Information Necessary

Background

The Centers for Disease Control and Prevention (CDC) is requesting OMB approval to reinstate the 'National Program of Cancer Registries Cancer Surveillance System (NPCR CSS), with changes (OMB Control No. 0920-0469, exp. 11/30/2012). The changes include 1) the addition of a preliminary data submission, 2) data definitions will be adjusted to maintain consistency with the American Joint Committee on Cancer terminology and definitions that reflect current medical practice (http://www.cancerstaging.org/), and 3) the reporting of additional data items by 10 registries funded through ARRA to support patient centered outcomes research (PCOR).

The NPCR CSS is an electronic data management system that produces comprehensive national statistics on cancer by compiling information collected under the authority of individual states, through central cancer registries located in 45 states, the District of Columbia, Puerto Rico, and the U.S. Pacific Island Jurisdictions; the central cancer registries in the remaining 5 states are funded by the National Cancer Institute (NCI). Data are combined from CDC and NCI to produce National cancer statistics. Information from the central cancer registries that are funded by CDC has been reported annually in prior years and each respondent receives a feedback report on its annual submission. For the period of this reinstatement request, the frequency of reporting will be changed from annual to semi-annual.

In the early 1990's, the NCI Surveillance Epidemiology and End Results Program supported and coordinated cancer registries in 5 states and 6 metropolitan areas covering 14% of the U.S. Population. Of the remaining states, 35 maintained some type of cancer registry though not all were complete and population based, and 10 states had no cancer registry beyond those maintained by large hospitals. 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 1a), later incorporated into the Public Health Service (PHS) Act [42 U.S.C. 242k]. This legislation authorizes the CDC to provide funds to states and territories to: 1) improve existing cancer registries; 2) plan and support infrastructure for 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. Information is collected and maintained at CDC under Section 301 of the Public Health Service (PHS) Act [42 U.S.C. 242k] (Attachment 1b).

State laws in all 50 states require the reporting of cancer cases diagnosed or treated within the state to the respective central cancer registry. The same is true for the District of Columbia, the Pacific Islands Jurisdiction and Puerto Rico and all additional references to states include these areas as well. The information on cancer cases is collected under the authority of the state, which receives data from hospitals, pathology laboratories, and providers who provide cancer diagnosis or treatment. The central registries then aggregate the data and use the information for activities such as state and county cancer reports, implementation of screening interventions, and cancer cluster investigations. CDC provides technical direction and establishes standards for data quality and completeness to support a cohesive National cancer surveillance system. Data from 5 states are reported to the NCI through their SEER program and data are combined at CDC to provide national statistics.

In fiscal year 2012, CDC awarded \$37 million through cooperative agreements to assist in funding central cancer registry operations in 45 states, the District of Columbia, Puerto Rico, and the Pacific Island Jurisdictions. To receive NPCR funding states must maintain State funding for their registry that was in place at the time of the initial CDC funding and also provide some matching funds or in-kind support.

Since 2000, state and territorial cancer registries have been reporting cancer incidence data to CDC. Cancer registration within the United States involves multiple partners to provide consistent and accurate reporting across the U.S. CDC works with Federal and Non-governmental Organizations to facilitate the development and implementation of national data standards, which change over time to reflect changes in cancer diagnosis and treatment.

Cancer is the second leading cause of death in the United States, second only to heart disease. In 2009, the most recent year for which complete information is available, more than 550,000 people died of cancer and more than 1.5 million were diagnosed with cancer (1). In addition to the personal impact of cancer, the financial burden is also substantial. The direct treatment costs of cancer in 2009 have been estimated at \$99.0 billion, with additional indirect costs of \$144.4 billion in lost productivity due to illness and premature death (2). It is estimated that 11.7 million Americans are currently alive with a history of cancer (3).

The cancer continuum framework spans prevention, early detection, diagnosis, treatment, survivorship and end-of life. There are several effective prevention and early detection measures that could substantially reduce the number of new cancer cases and prevent many cancer-related complications or deaths. To reduce the nation's cancer burden, behavioral and environmental factors that increase cancer risk must be reduced, and high-quality screening services and evidence-based treatments must be available and accessible, particularly to medically underserved populations (4, 5). The availability of complete and accurate cancer data at the National, state, and local level facilitates identification of disparities in cancer incidence and treatment, which supports targeted interventions such as screening or education to reduce the morbidity and mortality from cancer. The impact of prevention and early detection measures can also be monitored and

measured through state cancer registries since they are designed to monitor cancer trends over time, determine cancer patterns in various populations, and guide planning and evaluation of cancer control programs.

CDC's Division of Cancer Prevention and Control (DCPC) received \$20 million of the appropriated patient-centered health research funds from the American Recovery and Reinvestment Act (ARRA) of 2009, Public Law 111-5 (**Attachment 1c**). Ten central cancer registries received funding to harmonize their procedures and standards for managing more in-depth clinical data for colon, rectal, and breast cancers and chronic myeloid leukemia for cases diagnosed within their state in the year 2011. These data enhancements include clinically relevant biomarkers and specific treatment regimens, as well as co-morbid conditions that may affect treatment decisions. CDC provides technical direction and reporting standards for this effort.

Privacy Impact Assessment (PIA)

The NPCR CSS was certified in January 2011 under the Moderate Impact - Enterprise Master System Security Plan (EMSSP).

Overview of the Data Collection System

NPCR-funded grantees collect and aggregate cancer data under the authority of state laws which require cancer reporting to central cancer registries. Central cancer registries use these data for local public health purposes, and retain primary responsibility for information collection procedures.

In addition to work that is done within state boundaries, central cancer registries share cases through inter-state data exchange agreements to obtain cancer data on residents who travel to other states for diagnosis or treatment. Once quality control standards are met and the data are complete, they are ready for use and dissemination by the state and are also reported to CDC. This standard process conducted by all central cancer registries usually takes 12 to 24 months after the close of the year in which the cancer is diagnosed. After the data are reported to CDC from the individual central cancer registries, they are processed and data evaluation reports are generated.

For the 10 central cancer registries that receive ARRA funding, the same data collection process outlined above is followed. The main differences include:

- 1) the specificity of the treatment data variables that will be reported to CDC. For example, rather than reporting only that chemotherapy was given, the name of the actual chemotherapy agent will be provided; and
- 2) additional data linkages the state will conduct with other state data files such as hospital discharge data to identify co-morbid conditions which will be reported in standard NAACCR variables.

Information will be reported to CDC twice per year. CDC is requesting that once a year, NPCR registries report cumulative data which includes data from 1995 (for most registries) going forward. These data are considered final for reporting cancer statistics and are often referred to as *24-month data* since they are reported to CDC about 2 years after the year of diagnosis.

CDC requests to add a second data report, considered preliminary, that would involve earlier reporting of the most recent year of data. This additional data submission would be 1 year of data only. These data are referred to as *12- month data* since they will be reported to CDC approximately 1 year after the end of the diagnosis year.

The variables to be reported to CDC do not vary between the preliminary and final data submissions.

Data Report	Data Reported	Number of Respondents	Years of Data Reported
	Standard data variables	38	1 year of data
Preliminary (12 month)	Enhanced data variables	10	1 year of data
Final (24 month)	Standard data variables	38	Cumulative – all years from date of first registry reporting
	Enhanced data variables	10	Cumulative – all years from date of first registry reporting

Requested Submission Process

Items of Information to be Collected

The data items reported are based upon the North American Association of Central Cancer Registries (NAACCR) Standards for Cancer Registries, Volume II, which is a comprehensive reference to ensure uniform data collection, to reduce the need for redundant coding and data recording between agencies, and to facilitate the collection of comparable data among groups.

Thirty eight registries will report the standard list of data items and 10 ARRA funded registries will report an enhanced set of data items which includes variables that will support patient centered outcomes research

Attachment 3a is a copy of the submission specifications that were sent to NPCR grantees in August 2012 providing instructions for the reporting of Standard Final (24 month) cancer incidence data to CDC in November 2012. **Attachment 3a** also contains a list of data items for each of the two planned data submissions – preliminary and final. This table is updated annually based upon any changes outlined in the NAACCR Standards for Cancer Registries, Volume II.

The 10 ARRA funded registries will report an enhanced set of data items during each data submission. Not all of these additional variables are NAACCR variables. **Attachment 3b** is a copy of the document that contains enhanced data items that will be reported by the ARRA-funded central cancer registries. Some of the variables in **Attachment 3b** are IIF. For example, the biometric information such as height and weight and other medical information such as co-morbid conditions, specific treatment regimes, and biomarkers are considered IIF.

Information in identifiable form (IIF) is reported to CDC through NPCR CSS. Specifically, date of birth and medical information about the types of cancer that occur (histology, morphology, and behavior), the anatomic location, the extent of disease at the time of diagnosis, the kinds of treatment received by cancer patients, and the outcomes of treatment and clinical management are collected.

Identification of Website(s) and Website Content Directed at Children Under 13 Years of Age

Data are compiled at the state central cancer registry, are transmitted to NPCR CSS via a secure Website and are encrypted during transmission. The encryption is accomplished via Secure Sockets Layer (SSL) strong encryption, the same level of protection used by e-commerce sites to protect financial transactions. This Website for data collection has no content directed at children less than 13 years of age.

A2. Purpose and Use of the Information Collection

The NPCR 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.

Grantees have been 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 #DP12-12-1205), grantees report their cancer incidence data to CDC for these purposes and for public health and research analyses.

Under state law, central cancer registries routinely receive a standard set of data items on all cancer patients diagnosed in the state from hospitals, pathology labs, clinics and

private physicians. The cancer registries maintain these data items permanently in longitudinal databases that are used for public health surveillance, program planning and evaluation, and research.

The data enhancements from ARRA funding will allow state, Federal, University and other researchers to explore compliance with treatment recommendations for the selected cancers and short and long-term outcomes by treatment received. The comparison of different interventions and strategies to treat cancer would inform patients, providers, and decision-makers about those that are the most effective. The changes to the data collection instrument at the 10 ARRA funded central cancer registries impact the data reporting requirements by expanding the number of variables to be reported to CDC for breast, colon, and rectum cancers and chronic myeloid leukemia diagnosed in 2011. In addition, a few enhanced variables such as co-morbid conditions will be reported for all cancers within these 10 states. Given that the data submission process is automated and is the same irrespective of the number of variables submitted, the burden estimate is the same for the standard NPCR CSS report and the enhanced NPCR CSS report.

Not all central cancer registries receive ARRA funds. However, all central cancer registries that were funded by CDC through cooperative agreements were eligible to apply for the ARRA funds. After the ARRA funding has ceased, the 10 ARRA funded registries will continue to update data on vital status of the 2011 cases and report updated data through the normal NPCR data reporting process. The data will be analyzed by researchers within each state to assess patient centered outcomes and a data set will be established at the National Center for Health Statistics Research Data Centers to support research of a combined dataset by all qualified researchers.

The CDC and the states face the challenge of reducing cancer morbidity and mortality through prevention and early detection. Within CDC, the 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, including patient-centered health research with the receipt of ARRA funds.

CDC plans to add a second data submission, considered preliminary that would involve earlier reporting of the most recent year of data. The intent of this submission is to assess the completeness and quality of the most recent data and provide technical assistance as needed to improve data quality. As electronic reporting increases across the country, complete cancer incidence data should be available sooner. It is desirable to report cancer data more quickly since under the current process 2010 data will not be available for reports until 2013. Without the ability to assess and improve the preliminary data, CDC will not be able to make progress towards this long-term goal.

A data contractor, Inner City Fund (ICF) International, (Macro), has been retained to assist with data management and analysis. Based on annual CSS submissions, standardized reports which assess data quality and completeness are generated by ICF

Macro for the grantees and the CDC. These reports are used to guide additional technical assistance to the central registries as required. The performance indicators are also used for reporting to CDC officials, Congress and other national stakeholders.

When standards of completeness and quality have been met, CDC aggregates the data and uses it for the following purposes:

Cancer Surveillance: The CDC and the states face the challenge of reducing cancer morbidity and mortality through prevention and early detection. Effective cancer control requires the regular, ongoing collection and analysis of health-related data to monitor the frequency and distribution the disease in the population. The NPCR CSS helps 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 are enhanced by the ability to target areas with high rates of cancer screening. The Agency for Healthcare Research and Quality (AHRQ) (http://www.ahrq.gov/) includes measures for effectiveness of care in cancer. The AHRQ Healthcare Quality Report includes rates of advanced stage female breast and colorectal cancer and all invasive cervical cancer by state.

Since 2002, CDC and the NCI, in collaboration with NAACCR have published *United States Cancer Statistics (USCS) (http://www.cdc.gov/cancer/uscs)*. The *USCS* report contains a set of official federal cancer incidence statistics from each state that had high quality registry data. For cancer cases diagnosed in 2009, the most recent year for which federal data is available, 49 statewide population based cancer registries and District of Columbia met USCS publication criteria; however, data from Wisconsin for 2009 were suppressed at that state's request resulting in 98% population coverage. Data for selected cancer sites are also available as pre-calculated counts and rates on the NCI/CDC State Cancer 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) (9). The NPCR CSS continues to work to make timely data available for the NPHSS and publication in the *Morbidity and Mortality 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 CSS are 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*, the registries provide geographic coverage for 96% of the U.S. population. 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 created (**Attachment 4**) to provide a statistical basis for analyzing the cancer burden on a regional and national level (<u>http://www.cdc.gov/cancer/npcr/datarelease.htm</u>).

Given the stable infrastructure of the NPCR CSS, CDC is well positioned to contribute to patient-centered health research. With the data enhancements from ARRA funds, the NPCR CSS will be able to provide data for patient centered health research particularly focused on breast, colon, and rectum cancer and chronic myeloid leukemia.

Privacy Impact Assessment Information

Cancer registration is the fundamental method in the United States by which information is systematically collected about the occurrence of cancer, about the types of cancer that occur (histology, morphology, and behavior), the anatomic location, the extent of disease at the time of diagnosis, the kinds of treatment received by cancer patients, and the outcomes of treatment and clinical management. With the specificity of the medical information that is collected and the geographic coverage, NPCR CSS is able to derive more accurate and stable estimates of cancer incidence for population groups including racial and ethnic minorities, medically underserved groups, and other subpopulations; to conduct regional and national analyses to more accurately identify geographic variability in cancer treatment practices; and to promote greater access to cancer data for the general public, scientists, and policymakers. In addition, the data enhancements from ARRA funding would allow researchers to address the diagnosis and treatment of cancer using cancer registry data and other comparative effectiveness research questions. Prior to the use of data for cancer surveillance, program planning and evaluation, or research, data standards for completeness and quality must be met. Date of birth is required of all central cancer registries because some of the computerized edits used to check the data for its quality are written using the entire date of birth. Computerized edits are also used for the data items associated with the occurrence of cancer, about the types of cancer that occur (histology, morphology, and behavior), the anatomic location, the extent of disease at the time of diagnosis, the kinds of treatment received by cancer patients, and the outcomes of treatment and clinical management. For the central cancer registries that receive ARRA funds, height and weight will be collected and reported to help determine if the appropriate doses of chemotherapy or radiation were provided to the patient. For 2011 breast, colon, and rectal cancers and chronic myeloid leukemia, other medical information such as co-morbid conditions, specific treatment regimes, and biomarkers, will be collected to allow for appropriate comparison of the different interventions and strategies to treat cancer.

When the data have met the standards for completeness and quality, the data can be used. **Attachment 4** outlines two types of data sets: public-use data sets (PUDS) and restricted-access data sets (RADS). Current users of the NPCR CSS data must sign a data release agreement as outlined in **Attachment 4**, which is updated annually.

PUDS are defined as data sets that are comprised of aggregated data (i.e., not individual case-specific data) that have been modified as needed, according to accepted procedures, to block breaches of confidentiality and prevent disclosure of the patients' confidential information (10-16). PUDS do not contain information that is identifiable or potentially identifiable according to currently accepted procedures for reducing disclosure risk (10-16).

RADS are defined as versions of the full NPCR-CSS analytic data set (i.e., individual case-specific data) that have been modified as needed to minimize (but may not remove entirely) the potential for disclosure of confidential information. RADS do not contain personal identifiers such as a patient's name, street address, or social security number as this information is not transmitted by central cancer registries to CDC as part of their annual data submission. However, they may contain information that is potentially identifiable especially when linked with other data sets, such as the occurrence of a rare cancer in a person of a certain age or racial or ethnic group. Only the month and year (and not the full date of birth) are provided in this data set. Because restricted-access data sets may contain identifiable information, states have the option to exclude their data included in RADS.

DCPC has worked with the National Center for Health Statistics Research Data Center (NCHS RDC) to host the NPCR CSS RADS and the RADS for patient-centered health research. The NCHS RDC will allow researchers outside of CDC access to both of these RADS in a secure environment without jeopardizing the confidentiality of the data. The NPCR CSS RADS became available at the NCHS RDC in 2012 and the RADS for patient-centered health research will be available in 2013. NCHS will house the data securely and provide analytic support for data analyses. The NCHS RDC has an

established process for reviewing data proposals to access data. Once access has been approved and a researcher completes the analysis, NCHS RDC staff members review the data tabulations to ensure that confidentiality is maintained.

Since the information received by CDC or its contractor as part of NPCR CSS could lead to direct or indirect identification of cancer patients, CDC applied for and received 308(d) confidentiality protection approval in May 2000, with renewal in 2010 (see **Attachment 5**). In addition, any data published from NPCR CSS in surveillance reports, either in printed copy or on the Internet, are scrutinized to assure that the confidentiality of the individual is protected.

A3. Use of Improved Information Technology and Burden Reduction

The central cancer registries send their data to the CDC electronically via a secure socket layer (SSL) encryption using standard data definitions and record layouts. Central cancer registries complete the accompanying data submission forms electronically.

NAACCR standards help reduce errors and the electronic transmission of data is efficient and minimizes the reporting burden on the states (17).

A4. Efforts to Identify Duplication and Use of Similar Information

Cancer incidence data are available through the NCI's SEER Program, which represents 9%-26% of the population of the U.S. (<u>http://seer.cancer.gov</u>). 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 also not useful for the 45 states without a SEER registry for program planning and evaluation. NPCR-funded registries cover 96% of the United States population and complement the SEER data to provide 100% coverage of the U.S. population. In the states where the SEER program covers a part of the state (Alaska, California, 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. Four additional states (California, Kentucky, Louisiana, and New Jersey) receive funding from NCI beginning in 2002 to enhance the representativeness of the SEER program.

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) (18). 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 1,400 participating hospitals. The program was started in 1989 and approximately 75% of all U.S. cancer cases 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. In addition, the NCDB cannot identify multiple cancer reports for the same individual that may arise from a surgical hospital, pathology lab, and treatment facilities. This means that a single case may be reported multiple times in the database.

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. With the addition of data items that will be collected by the central cancer registries that receive ARRA funding, the NPCR CSS will be able to address the diagnosis and treatment components of the cancer continuum using cancer registry data for breast, colon, and rectum cancer and for Chronic Myeloid Leukemia. The comparison of different interventions and strategies to diagnose and treat cancer would inform patients, providers, and decision-makers about those that are the most effective.

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 of final data (24-month data) occurs on an annual basis in November 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. In addition, the ability to monitor trends in cancer incidence would be jeopardized. It is essential that CDC and State program managers evaluate program strengths and weaknesses on an annual basis and make adjustments. Because of staff changes or other issues, a central registry that has performed well in the past may have problems with data quality or completeness. It is critical to identify these registries as soon as possible so that needed technical assistance or guidance can be provided. 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. The data are also used to evaluate the success over time of prevention (tobacco control) or screening (breast, colorectal, cervical) efforts at the state and National level. The report of the enhanced variables will allow for patient centered outcomes research, which could affect treatment recommendations and assessments of treatment efficacy.

During the period of this reinstatement, CDC will begin requesting two submissions per year. The additional preliminary submission in the winter or spring will allow CDC to

prepare to produce early estimates of cancer incidence and other statistics, in advance of the detailed data validation steps required for the fall submission. Also, the assessment of data quality will allow for earlier technical assistance to central registries which are having difficulties. As electronic reporting continues to grow, this earlier submission provides a foundation for reporting final data earlier for public health purposes. Without this initial work to evaluate and improve the quality of this earlier data, progress cannot be made towards this goal.

There are no legal obstacles to reduce the burden.

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

This request fully complies with the regulation 5 CFR 1320.5.

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 December 10, 2012, in Volume 77, No. 237 pp. 73469-73470 (Attachment 6). No public comments were received from the 60-Day Federal Register Notice.

B. Attachment 7 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 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.

Since the inception of NPCR CSS, CDC staff receive continuous feedback from grantees on the annual data submission packet **(Attachment 3a)** and the data release policy **(Attachment 4)** through regular scheduled conference calls of the NPCR Central Cancer Registry Council and the NPCR Scientific Working Group. The two groups are facilitated by CDC and workgroup members include a subset of the NPCR grantees.

For the data that is enhanced by ARRA funding, CDC staff will continue to receive continuous feedback from central cancer registries since ICF Macro will provide

technical guidance on best practices and will facilitate information sharing across the registries through e-mail and regularly scheduled conference calls.

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

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, with renewal in 2010 **(Attachment 5)**. The proposed new data collection will be covered under an extension of the 308(d). This approval was updated and approved November 10, 2010, to include the specific data items that will be collected by the central cancer registries that receive ARRA funding.

The risk of direct identification of an individual in NPCR CSS data is remote because personal identifying data (name, social security number and street address) are not reported to the CDC. However, a unique identifier assigned by the state to each individual cancer patient is 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 identities 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. However, that unique identifier is assigned by the NPCR CSS contractor, ICF Macro, and is not the same identifier reported by grantees.

Of greater concern is the geographic data (e.g., county, census tract, zip code) that are 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 via the National Center for Health Statistics Research Data Center (NCHS RDC), 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 4). In the first tier of data (the least confidential), state would be the smallest geographic unit released. A 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 month and 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. 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.

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 are used (19). NCHS has guidelines for published data and one for microdata 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 than five unweighted cases", and 3) "In no case should a quantity figure be based upon fewer than five unweighted 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 serve as a model for NPCR CSS as confidentiality procedures. In addition, the program will need to be attentive to changes in the environment that may impact efforts to maintain confidentiality.

Collection of incidence data for surveillance is considered public health surveillance rather than research. However, epidemiologic analyses of the NPCR data for research is covered by a protocol submitted to Institutional Review at CDC. The study protocol (#2594) for analysis of data from NPCR-CSS has been reviewed and approved by a CDC Institutional Review Board (IRB). The most current notice of approval (February 8, 2013) is attached **(Attachment 8).** The Division of Cancer Prevention and Control maintains IRB approval through the annual continuation process. Analysis of the enhanced data collected through the ARRA funding was incorporated into the IRB protocol for NPCR-CSS.

Privacy Impact Assessment Information

A. This submission has been reviewed by CDC which determined that the Privacy Act does not apply. 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.

B. The NPCR CSS data are secured by technical, physical, and administrative safeguards. A data contractor, ICF Macro, in Bethesda, Maryland, has been retained to assist with data management and analysis. The safeguards are outlined below:

Technical

- The NPCR CSS project is undergoing the required Security Certification and Accreditation renewal process managed by CDC's Chief Information Security Officer.
- The NPCR CSS project data reside on a dedicated server that resides on ICF Macro's local area network (LAN) behind the contractor's firewall and is password protected on its own security domain. Access to the NPCR CSS server is limited to the contractor's authorized project staff. No other non-project staff are allowed access to the NPCR CSS. All of the contractor's project staff are required to 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 the data collection contractor'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.
- Once the data have been compiled by the contractor and delivered to CDC via the document server behind the firewall, all NPCR CSS datasets are maintained for restricted access on CDC's secure LAN server.

Physical

- The contractor's NPCR CSS server is housed in a secure, guarded facility. All contractor staff are issued identification badges. Elevator and stairwell access is controlled by key cards.
- 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 are stored under lock and key when not in use and will be destroyed when no longer needed.
- Once the data have been compiled by the data collection contractor and delivered to CDC, all NPCR CSS datasets are maintained for restricted access on a secure LAN server, which is housed in a secure facility. All CDC staff are issued identification badges and access to the building is controlled by key cards.

Administrative

• CDC staff and the contract staff have developed a security plan to ensure that the data are kept secure and confidential. Periodic review and update of the data collection contractor's 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.

- All project staff receive annual security awareness training covering security procedures. The contractor's project security team oversees operations to prevent unauthorized disclosure of the NPCR CSS data.
- Once the data have been delivered to CDC, access to these datasets is only granted when appropriate confidentiality release forms have been signed and returned to the NPCR CSS Data Security Steward.

C. The respondents for the NPCR CSS are central cancer registries, not individuals. Each central cancer registry is responsible for working with local sources of cancer information to comply with applicable local requirements relevant to informing patients about the intended uses of the information collection and any plans for sharing the information.

D. NPCR-funded central cancer registries report patient-level information to CDC twice annually. In addition, a subset of 10 central cancer registries that receive ARRA funding will report additional patient-level information to CDC with each NPCR CSS submission. 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, with renewal in 2010 covering the additional data reported by the central cancer registries that receive ARRA funding **(Attachment 5)**.

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) are not reported to the CDC. However, a unique identifier assigned by the state to each individual cancer patient is 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. However, that unique identifier is assigned by the NPCR CSS contractor, ICF Macro, and is not the same identifier reported by grantees.

A11. Justification for Sensitive Questions

There are no sensitive questions asked directly of the central cancer registry (i.e., the respondent); however, some of the patient-level data received from the central cancer registry are of a sensitive nature. This data include 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. The information is required to meet cancer surveillance objectives.

Information regarding weight, height, co-morbid conditions, specific treatment regimens, and biomarkers are central to patient centered outcomes research. The information is required to allow for appropriate comparison of the different interventions and strategies to diagnose and treat cancer.

A12. Estimation of Annualized Burden Hours and Costs

A. Respondents are the 48 NPCR central cancer registries (45 states, the District of Columbia, Puerto Rico, and the Pacific Islands Jurisdiction). CDC is requesting two submissions. Each of these submissions will include 38 registries reporting standard data items listed in **Attachment 3a**. Ten ARRA funded central cancer registries will report the same variables plus additional data items that will support patient centered outcomes research (variables are listed in **Attachment 3b**).

The estimated burden per response is 2 hours. All information is reported to CDC electronically. The total estimated annualized burden is 192 hours. 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 9).** This information appears on the log in page of the website that the states use to transfer their files electronically.

Type of Respondents	Form Name	No. of Respondents	No. of Responses per Respondent	Average Burden per Response (in hours)	Total Burden (in hours)
Central Cancer Registries in	Standard NPCR CSS Report	38	2	2	152
States, Territories, and the District of Columbia	Enhanced NPCR CSS Report	10	2	2	40
				Total	192

Table A12-A. Estimated Annualized Burden Hours

B. The annualized cost to respondents of reporting data to CDC is estimated to be \$6,912. It is estimated that the following state cancer registry personnel will be required to help prepare and submit data electronically to CDC: data managers, information technology staff, and program directors. However, it should be noted that the specific nature of the work in the central cancer registries does not correlate with the employment categories as outlined by the Department of Labor. The categories listed below are similar in job description to those in central cancer registries.

Type of Respondents	Form Name	No. of Respondents	Total Burden (in hours)	Average Hourly Wage	Total Cost
Central Cancer Registries in States,	Standard NPCR CSS Report	38	152	\$36	\$5,472
Territories, and the District of Columbia	Enhanced NPCR CSS Report	10	40	\$36	\$1,440
				Total	\$6,912

Table A12-B. Annualized Cost to Respondents

*Based upon U.S. Bureau of Labor Statistics. *Occupational Employment Statistics. May 2009 National Occupational Employment and Wage Estimates.* Washington, DC: U.S. Bureau of Labor Statistics. Available at: http://www.bls.gov/oes/current/oes_nat.htm#00-0000 [accessed August 23, 2010.]

A13. Estimates of Other Total Annual Cost Burden to Respondents or 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 NPCR CSS data collection is \$1,600,955 per year for a five-year total of \$8,004,777. A data management contract was awarded to ICF Macro in calendar year 2008. Additional annual costs include personnel costs of federal employees involved in oversight and analysis. The annual staff cost is estimated at \$120,000 (1.0 epidemiologist FTE, 0.2 public health advisor FTE, and miscellaneous expenses include travel, etc.).

|--|

A15. Explanation for Program Changes or Adjustments

This Reinstatement request includes an increase in burden attributable to a change in frequency from annual reporting to semi-annual reporting. The previous information

collection request was based on an annualized estimate of 96 hours (48 respondents x 1 response/year x 2 hours/response = 96 hours). In this Reinstatement request, annualized burden is estimated to be 192 hours (48 respondents x 2 responses/year x 2 hours/response = 192 hours).

This Reinstatement request divides the 48 respondents (central cancer registries) into two groups: a group of 38 respondents that will submit a standard report (Attachment 3a), and a group of 10 respondents that will submit an enhanced report which includes additional variables (Attachment 3b). However, since the data files are electronically created, there is no additional burden associated with including the additional variables in the electronic data file (report). Therefore the estimated burden per response has not changed and is the same for both groups of respondents. For all respondents, the estimated burden per response is 2 hours.

A16. Plans for Tabulations and Publication and Project Time Schedule

CDC is requesting that preliminary data (referred to as 12 month data) be submitted in summer (winter beginning in 2014). These data will consist of one year of data for the most recent year of available cancer data. These data will be evaluated for completeness and quality and reports will be provided back to each registry and to CDC. Additional technical assistance will be provided as needed.

CDC is requesting that NPCR registries report Final Data (24-month data) each year in November. These data include diagnoses for each year the registry has been funded by NPCR to report data (1995 forward in many cases). Corrections and additions are added each year by the central cancer registries. Consequently 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 similar to what is found here:

Table A16. Time Schedule for	Preliminary and Final Data Report	ing, Analysis and
Publication		

Summer 2013	
(Winter thereafter)	
September 2013	

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 Submissions

There are no exceptions to the certification.

REFERENCES

- 1 U.S. Cancer Statistics Working Group. *United States Cancer Statistics:* 1999–2009 *Incidence and Mortality Web-based Report.* Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2013. Available at: <u>www.cdc.gov/uscs</u>.
- 2 National Heart, Lung, and Blood Institute. *NHLBI Fact Book, Fiscal Year 2008*. Bethesda (MD): National Heart, Lung, and Blood Institute; 2009.
- 3 Howlader N, Noone AM, Krapcho M, Neyman N, Aminou R, Altekruse SF, Kosary CL, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Eisner MP, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2009 (Vintage 2009 Populations), National Cancer Institute. Bethesda, MD, <u>http://seer.cancer.gov/csr/1975_2009_pops09/</u>, based on November 2011 SEER data submission, posted to the SEER web site, April 2012.
- 4 Curry SJ, Byers T, Hewitt M. *Fulfilling the Potential of Cancer Prevention and Control*. Washington (DC): The National Academies Press; 2003.
- 5 Haynes MA, Smedley BD. *The Unequal Burden of Cancer: An Assessment of NIH Research and Programs for Ethnic Minorities and the Medically Underserved.* Washington (DC): The National Academies Press; 1999.
- 6 Tucker TC, Howe HL, Weir HK. Certification of population-based cancer registries. *Journal of Registry Management* 1999; 26(1): 24-27.
- 7 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.
- 8 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.
- 9 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.
- 10 Centers for Disease Control and Prevention. *CDC/ATSDR/CSTE Data Release Guidelines for Re-Release of State Data*. Atlanta: Centers for Disease Control and Prevention; 2003 (available upon request).
- 11 American Statistical Association. Privacy, Confidentiality, and Data Security Web Site. Alexandria, VA: American Statistical Association; 2005. Available at <u>http://www.amstat.org/comm/cmtepc/index.cfm</u>.
- 12 Doyle P, Lane JI, Theeuwes JM, Zayatz LM (eds). *Confidentiality, Disclosure, and Data Access: Theory and Practical Application for Statistical Agencies*. Amsterdam: Elsevier Science BV; 2001.
- 13 Federal Committee on Statistical Methodology. *Checklist on Disclosure Potential of Proposed Data Releases.* Available at <u>http://www.fcsm.gov/committees/cdac/cdac.html</u>.
- 14 Federal Committee on Statistical Methodology. *Report on Statistical Disclosure Limitation Methodology*. (*Statistical Working Paper* 22). Washington (DC): Office of Management and Budget; 1994. Available at <u>http://www.fcsm.gov/working-papers/spwp22.html</u>.

- 15 McLaughlin C. Confidentiality protection in publicly released central cancer registry data. *Journal of Registry Management* 2002; 29(3):84-88.
- 16 Stoto M. *Statistical Issues in Interactive Web-Based Public Health Data Dissemination Systems.* Draft report prepared for the National Association of Public Health Statistics and Information Systems, Rand Corporation; September 2002.
- 17 Havener, L and Thornton, M (editors). Standards for Cancer Registries, Vol II: Data Standards and Data Dictionary, 13th ed. Springfield, Ill: North American Association of Central for Cancer Registries; 2008.
- 18 North American Association of Central Cancer Registries. *Cancer in North America*, 2000-2006: volumes 1-3. Springfield (IL): North American Association of Central Cancer Registries; 2009. Available at : <u>http://www.naaccr.org/index.asp?</u> <u>Col SectionKey=11&Col ContentID=50</u>
- 19 National Center for Health Statistics. *NCHS Staff Manual on Confidentiality*. Hyattsville, MD: Centers for Disease Control and Prevention, National Center for Health Statistics; 1997.