ATTACHMENT 11

Supplemental Data Collection and Privacy Information

Supplemental Grantee Data Collection and Privacy Information

Overview of the Data Collection System

NBCCEDP-funded grantees, which are state, territorial and tribal governments or bona-fide agents, collect data to manage their screening programs and retain primary responsibility for information collection procedures. A subset of clinical data collected by grantees is reported to CDC to describe the services provided through the program.

Grantee programs establish a provider network for cancer screening and diagnostic services and collect clinical data from these providers on each client screened through the program. Each grantee program is responsible for developing an appropriate consent form, and for implementing assurances that all sensitive and/or personally identifiable information collected is properly maintained and secured. Grantees maintain data in local data management systems used to administer their programs, reimburse providers for services, ensure quality of services, and track patients for appointments and follow-up. Grantees can use any software system that meets their needs. They have the option to use a software application provided by CDC, a state health department based data system, or a data system customized to meet local needs. The optional software provided by CDC is a Windows-based desktop application that supports patient tracking and facilitates the extraction of data to report to CDC. CDC provides any necessary technical support to grantees that use this data management system.

Grantees are required to establish a Memorandum of Agreement with their corresponding state Central Cancer Registry (CCR) and link records of cancer cases diagnosed through their screening programs for case reporting and quality assurance. State CCRs are the primary source of end result information collected on cancer cases, and the linkages are used by NBCCEDP grantees to confirm diagnostic outcomes and collect a discreet set of standardized registry data on cancer stage at diagnosis, a measure of the extent/spread of disease and a critical indicator to evaluate the effectiveness of a cancer screening program.

Twice a year, grantees aggregate and report a subset of their patient-level clinical data to CDC to monitor and evaluate the program. Prior to electronic data transfer to the data contractor, each grantee removes all direct personal identifiers (such as name) and assigns a unique code for each client in the data base. The CDC will not accept a method of record identification, such as social security number, that may be linked to other databases. The development of a unique method of record encryption and

identification by each grantee program permits the CDC to anonymously track each client served throughout their association with the NBCCEDP, without the use of names. The grantees maintain the encryption information between their unique codes and the personal identifiers in their database. Neither the encryption scheme nor identifying information on the client, other than the fields noted, is ever provided to the CDC or the data contractor.

Grantees submit the MDE data as an electronic fixed-length text file using a secure submission web site (Attachment 3b), which simplifies the data reporting process for grantees and organizes the receipt of grantee text files by the CDC. Information is archived indefinitely. The contractor aggregates and validates the data for quality and completeness and prepares a SAS analysis file and a set of feedback reports to CDC and grantees within 60 days of the submission. The analysis file contains the same patient ID code that is submitted by the grantees and the day of birth is recoded to equal '15' prior to submission to CDC. Once data have been compiled by the contractor and delivered by courier service to CDC, all NBCCEDP datasets are maintained for restricted access on CDC's secure LAN server.

Items of information to be collected

Twice a year, grantees aggregate and report a subset of their patient-level clinical data to CDC to monitor and evaluate the program. The data submission includes cumulative records since the inception of the grantee screening program through a cutoff-period that allows a minimum of 3.5 months to collect, validate and prepare the data for submission. These data include coded patient identifiers, screening history, demographic data, screening and diagnostic tests provided and results, diagnostic outcomes, and treatment initiation information if cancer is diagnosed. The list of specific data items is provided in Attachment 3a (MDEs). Information in Identifiable Form (IIF) is collected. All screening records include date of birth and medical information on cervical and breast cancer screening and diagnostic tests and results. If cancer is detected, IIF is collected on the characteristics (histology, behavior, stage) of the cancer diagnosed and the dates the client was diagnosed and started treatment.

How information is shared and for what purpose

The clinical data is used by CDC to monitor and evaluate the NBCCEDP, ensure the quality of clinical services, provide feedback to grantees and Congress on program outcomes, evaluate the costs and effectiveness of the program, and inform program planning and policy decisions for organized cancer screening programs.

There is no public-use dataset of NBCCEDP clinical services. DCPC investigators have restricted access to an analytical dataset of program results to use for analysis and publication in peer-review journals and presentations to cancer control organizations. Program participation and results are be reported in aggregate to describe client demographics, volume of screening, cancer detection rates by screening test type, diagnostic outcomes by age, race/ethnicity, and geographic region, complication rates, and quality assurance of patient follow-up and adherence to cancer screening guidelines. .

Formal reports are developed for publication both semi-annually and periodically. These reports present results from the program based upon demographic information such as age and race, and reported as national aggregate data rather than grantee-specific. A limited set of grantee-specific data reported in a five-year aggregate are available to the public on the CDC web site. Reports do not include record identifiers. Reports are disseminated to the public through the CDC public web site, peer review journals, and publications. Any data for publication are scrutinized to assure that small cell counts are masked and the privacy of the individual is protected. Secondary analysis of data for the purposes of research is conducted to address specific research questions concerning breast and cervical cancer screening.

Investigators outside of the agency are permitted to submit a proposal to the CDC requesting use of the national data set. Each proposal is reviewed internally by a committee comprised of designated representatives from each Branch in the DCPC who are knowledgeable about the data set, its uses, and limitations. Upon review by the committee, each proposal is approved, denied, or the investigator is asked to provide additional information. Investigators who submit successful proposals to the CDC are required to sign a Data Sharing Agreement for Special Use Form (Attachment 6) indicating they agree to comply with the provisions outlined for data use and include a DCPC collaborator. Successful applicants do not gain access to the entire data set. The CDC develops and provides to each successful applicant a custom data set that meets the minimum needs of their proposal. The CDC replaces all encrypted identifiers in custom data sets with randomly generated record identifiers that cannot be linked back to the CDC database or to any of the identifying information maintained in the grantee databases.

<u>Statement of impact on the respondent's privacy</u>

The analysis dataset at CDC does not contain direct personal identifiers as this information is not available to CDC. As such, the data collection will have little or no effect on the respondent's privacy. However, it may contain information that is potentially identifiable especially when linked with other

datasets, such as in the occurrence of a cancer in a person of a certain combination of age, race, ethnicity and geographical information.

Opportunities to consent to sharing and submission of information

The respondents for the NBCCEDP are NBCCEDP grantees, not individuals. Each grantee program is responsible for collecting an appropriate consent form from clients before enrolling them in the program to receive screening services, and for developing procedures applicable within their jurisdiction to inform patients about the intended use of the information collection and any plans for sharing the information. Data are treated in a secure manner and will not be disclosed, unless otherwise compelled by law.

How information is secured

The MDE data are secured by technical, physical and administrative safeguards as outlined below.

Technical

- The MDE data reside on a dedicated server that resides on the contractor's local area network behind the contractor's firewall and is password protected on its own security domain. Access to the server is limited to the contractor's authorized project staff. No non-project staff is allowed access to the data. All of the contractor's project staff is required to sign a confidentiality agreement before passwords and keys are assigned.
- MDE data that are submitted electronically from grantees via a password-protected secure website (Attachment 3b) are encrypted during transmission and arrive on a server behind the data collection contractor's firewall. Each grantee has its own directory location so no grantee has access to another grantee's data. 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.
- Once data have been compiled by the contractor and delivered to CDC via courier, all MDE data are maintained for restricted access on CDC's secure LAN server.

Physical

- The contractor's server is housed in a secure facility with restricted access.
- Receipt and processing logs are maintained to document data receipt, file processing and report production. All reports and electronic storage media containing MDE data are stored under lock and key when not in use and will be destroyed when no longer needed.
- Once data have been compiled by the data contractor and delivered to CDC, all CCDE

datasets are maintained for restricted access on a secure LAN server, which is housed in a secure facility. All CDC staff is issued identification badges and access to the building is controlled by key cards.

Administrative

- CDC and contract staff have developed and implemented an information system security plan to ensure that the data are kept secure. Periodic review and update of the data contractor's security processes is conducted to adjust for needed changes and will be amended as needed to maintain the continued security of the data.
- The contractual agreement between CDC and the contractor includes non-disclosure terms. The contractor's project security team oversees operations to prevent unauthorized disclosure of the CCDE data.
- Once the data have been delivered to CDC, data are housed on CDC's secure LAN server and restricted access is controlled by the MDE data manager.