



**February 2, 2007**

**To: U.S. Department of Health and Human Services  
Room 434E  
200 Independence Avenue S.W.,  
Washington, D.C., 20201  
Attention: Personalized Health Care RFI**

**From: Dr. Eugenia E. Calle, PhD  
Managing Director, Analytic Epidemiology  
American Cancer Society  
1599 Clifton Road NE  
Atlanta, GA 30329  
404-329-5741  
jcalles@caner.org**

**Re: Response to U.S. Department of HHS Request for Information (RFI):  
Improving Health and Accelerating Personalized Health Care Through  
Health Information Technology and Genomic Information in Population-  
and Community-based Health Care Delivery Systems**

**Background -- HHS Secretary Michael Leavitt's 500-Day Plan**

A major component of Health and Human Services (HHS) Secretary Michael Leavitt's 500-Day Plan is to "Transform the Health Care System." His vision includes "health records that are linked through an interoperable system that protects privacy," and his strategy is to "advance health information technology by establishing standards for interoperability, privacy, and data exchange and developing prototypes for the National Health Information Network." The following proposal is for one such prototype, offered as an illustration of how existing data on cancer registration can be used to greater potential to create the "cancer backbone" of the electronic health record (EHR).

**PROPOSAL:**

Establish a nationwide database of all newly diagnosed cases of cancer in the U.S that is secure and electronically interoperable and that can be linked with the EHR. The database will include personal identifiers, allowing it to be linked to all other electronic databases. Two key features of this proposed database are centralization and inclusion of personal identifiers. Qualified researchers and practitioners wishing to use this database to evaluate or improve cancer patient care or for other cancer research would utilize a centralized application, review, and IRB-approval process. The proposed database would

serve as a national platform for information regarding cancer surveillance, cancer etiology, and cancer patient care. It would greatly facilitate identification of all new cases of cancer occurring in the U.S. This database would serve as the “cancer backbone” of the EHR. *Importantly, a great deal of the infrastructure needed to create such a centralized, standardized and searchable database already exists.*

#### **RATIONALE:**

Cancer is the only reportable chronic disease in the U.S. Virtually every state has legislation to establish a state-wide cancer registry. Although the Federal government publishes national data on cancer incidence (see below), cancer registration is largely a state responsibility. Cancer researchers conducting large national or multi-state studies must now apply to individual State cancer registries to ascertain new cases of cancer in their study populations. For each individual state registry, researchers must make an application, go through a local IRB review, submit their study population to the registry, monitor the linkage process, receive records of cancer cases from the registry, and review the results of the linkage for accuracy and completeness. This process varies for each registry, making it logistically unnecessarily complex. In the case of long term studies, this procedure is repeated on a regular basis.

#### **The proposed database will facilitate the following:**

1. Detailed information about every newly diagnosed case of cancer can rapidly be incorporated into the EHR through linkage with the proposed database. Information regarding cancer diagnoses, tumor types, histologic and pathologic characteristics of tumors, etc. is already being collected by state registries in a uniform fashion. Data standards for reporting of these data have already been developed and tested and are in current use.
2. Treatment data are collected by some registries and the incorporation of more complete treatment data into more registry records in a standardized manner is likely to occur in the next several years. As this effort moves forward, a centralized database will allow for evaluation of the quality of cancer patient care nation-wide and for identification of disparities in cancer patient care. In addition, a centralized database would allow for a nation-wide linkage of cancer occurrence with Medicare data. The existing SEER-Medicare database is currently the major source of data for studies of cancer patient care, but SEER registries cover only 26% of the U.S. population (see below).
3. An additional benefit will come from linkage of the proposed national cancer database to the existing National Death Index (NDI) to ascertain mortality among all cancer cases and determine how variability in quality of care influences survival. This will allow for nationwide surveillance of cancer survival and mortality.
4. Second primary cancers that occur as late effects of treatment among cancer survivors (childhood and adult) can be easily ascertained and tracked through this database.
5. New cases of cancer in large prospective follow-up studies of cancer etiology can be easily and rapidly identified in this database. Many large cohort studies forgo the complex attempts at linkage with state cancer registries and rely on other costly strategies such as verifying cancer cases reported by their study participants by obtaining medical and pathology records from hospitals all over the country. With no central database of cancer incidence (with identifiers), scientists are spending large sums of money, a great

deal of effort and considerable time completing the cancer case ascertainment needed for their research.

6. Selection of cancer cases for population case-control studies can be done nationwide rather than within a single or a few registries (which is the current model), allowing for easier study of rare cancers and examination of regional differences.

## **EXISTING RESOURCES AND OPPORTUNITIES:**

### **Federal Programs:**

Currently, information on newly diagnosed cancer cases in the U.S. is collected by registries in two federal programs: NCI's Surveillance, Epidemiology, and End Results (SEER) Program and CDC's National Program of Cancer Registries (NPCR).

The SEER program was established over 30 years ago and currently collects and publishes cancer incidence and survival data from 14 population-based cancer registries and 3 supplemental registries covering approximately 26% of the US population.

The NPCR was established in 1992 and funds 49 statewide and territorial cancer registries. NPCR registries currently cover 96% of the US population.

These two federal programs have already developed a fruitful partnership with each other and with the National Center for Health Statistics (NCHS) to assemble official federal statistics on cancer incidence and mortality for the entire US. This partnership has for the past four years produced the publication "United States Cancer Statistics (USCS)" as well as electronic databases underlying the publication.

### **Collaborating Partner of the Federal Programs:**

Both of the federally funded registry programs work closely with the North American Association of Central Cancer Registries (NAACCR) to promote cancer incidence surveillance in the United States (and Canada). NAACCR was established in 1987 as an umbrella organization for population-based cancer registries, governmental agencies, professional associations, and private groups in North America interested in improving the quality and use of cancer registry data. All NPCR and SEER registries are members of NAACCR. In 1997, NAACCR developed a certification process for recognizing registries that achieve high-quality data standards. NAACCR also receives and publishes aggregate data from most NPCR and SEER registries.

### **Opportunity to Leverage Existing Resources:**

All of the individual records of cancer occurrence in the United States are already being collected by one or both of these federal programs. The two programs already work together to aggregate the data for the entire U.S. *What is missing?* The individual records in the NPCR, USCS, and NAACCR databases contain no personal identifiers. There is no ability to electronically link these databases with any other population databases. (NCI's SEER database does include personal identifiers – however, external researchers cannot apply to NDI for linkage of their study participants to the entire SEER database). A centralized nationwide registry that can be linked to other study populations would greatly facilitate cancer research in this country.

## **PROOF OF PRINCIPLE: The National Death Index**

Many of the underlying principles and operational policies used to establish the National Death Index (NDI) have relevance and can serve as a model today for the proposed nationwide database of cancer incidence. The NDI was established in 1979 to simplify the procedure of determining whether or not an individual had died anywhere in the U.S. and if so, the location of his/her death record. In 1993, it was expanded to provide cause of death in addition to fact of death. Death data in NDI comes from each state under separate contractual agreements between the individual states and the NCHS. The NCHS is the processor and compiler of the records supplied by the states, and researchers are the consumers of the NDI data.

In 1977, prior to the establishment of the NDI, the Director of the NCHS established an Ad Hoc Working Group to Develop Plans and Procedures for the Possible Implementation of a National Death Index. The group began with the idea that the NDI would be strictly used as a research tool. They developed a final report in 1978 that included recommendations for general policies for implementing the NDI. Some key aspects of these policies included:

- Use of the NDI would require an application process and would only be for statistical research purposes
- Detailed criteria would be developed and applied for approval of NDI applications
- An NDI Advisory Panel would review and approve all applications and would be composed of representatives from States, NIH, NIOSH, NCHS, university, and industry.
- The NDI maintains individual contracts with the States and States are reimbursed on a fixed unit price cost
- Contracts with the States include a confidentiality clause reviewed by the Office of the General Counsel, Public Health Service, Dept of Health, Education and Welfare (now HHS)

## **CHALLENGES:**

1. State cancer registries must agree to share the personal identifiers that accompany their cancer records. To do this, they will need to be assured of many things, including:
  - they will not lose ownership of the data,
  - they will be reimbursed for this information,
  - the confidentiality of this information will not be breached,
  - such data sharing will be beneficial for them
    - by removing their current burden of research-related linkages to their data
    - by replacing any monies they are receiving for such linkages
  - their state laws will be accommodated
2. Effective collaboration and partnership between CDC, NCI, and NAACCR to accomplish this larger goal will be essential.
3. A federal agency must be chosen to compile and maintain this database and to oversee the application/review process. Example agencies that might fill this role: HHS, CDC, NCI, NCHS.
4. Currently the Veteran's Administration does not have a mandate to report cancer cases to state cancer registries. These cases comprise approximately 6-10% of male cancer cases nationwide, and more in some states.

**RECOMMENDATION:**

We recommend that Secretary Leavitt appoint an Ad Hoc Working Group to develop plans and procedures for the possible implementation of a National Cancer Registry that includes personal identifiers so that it is linkable to other electronic databases. This group should minimally include representatives from the CDC, NCI, NCHS, NAACCR, the States, and potential users of the registry e.g. the American Cancer Society (ACS) and university scientists with large geographically diverse study populations, and other appropriate stakeholders.