

Supporting Statement A for

Population Sciences Biospecimen Catalog (PSBC) (NCI)

September 8, 2015

Danielle Carrick  
Program Director  
Epidemiology and Genomics Research Program  
Division of Cancer Control and Population Sciences (DCCPS)  
National Cancer Institute (NCI)

9609 Medical Center Dr., 4E224  
Rockville, MD 20892

Phone: 240-276-6749  
Danielle.Carrick@nih.gov

Table of Contents

**A. JUSTIFICATION.....**

A.1 CIRCUMSTANCES MAKING THE COLLECTION OF INFORMATION NECESSARY.....

A.2. PURPOSE AND USE OF THE INFORMATION COLLECTION.....

A.3 USE OF INFORMATION TECHNOLOGY AND BURDEN REDUCTION.....

A.4 EFFORTS TO IDENTIFY DUPLICATION AND USE OF SIMILAR INFORMATION.....

A.5 IMPACT ON SMALL BUSINESSES OR OTHER SMALL ENTITIES.....

A.6 CONSEQUENCES OF COLLECTING THE INFORMATION LESS FREQUENTLY.....

A.7 SPECIAL CIRCUMSTANCES RELATING TO THE GUIDELINES OF 5 CFR 1320.5.....

A.8 COMMENTS IN RESPONSE TO THE FEDERAL REGISTER NOTICE AND EFFORTS TO CONSULT OUTSIDE AGENCY.....

A.9 EXPLANATION OF ANY PAYMENT OF GIFT TO RESPONDENTS.....

A.10 ASSURANCE OF CONFIDENTIALITY PROVIDED TO RESPONDENTS.....

A.11 JUSTIFICATION FOR SENSITIVE QUESTIONS.....

A.12 ESTIMATES OF HOUR BURDEN INCLUDING ANNUALIZED HOURLY COSTS.....

A.13 ESTIMATE OF OTHER TOTAL ANNUAL COST BURDEN TO RESPONDENTS OR RECORD KEEPERS.....

A.14 ANNUALIZED COST TO THE FEDERAL GOVERNMENT.....

A.15 EXPLANATION FOR PROGRAM CHANGES OR ADJUSTMENTS.....

A.16 PLANS FOR TABULATION AND PUBLICATION AND PROJECT TIME SCHEDULE.....

A.17 REASON(S) DISPLAY OF OMB EXPIRATION DATE IS INAPPROPRIATE.....

A.18 EXCEPTIONS TO CERTIFICATION FOR PAPERWORK REDUCTION ACT SUBMISSIONS.....

## **LIST OF ATTACHMENTS**

Attachment 1. Population Sciences Biospecimen Catalog

Attachment 2. Letter to PIs

Attachment 3. Annual Update Letter to PIs

## **A. Justification**

### **Abstract**

This is a request for approval of a new collection. The National Cancer Institute (NCI) Division of Cancer Control and Population Sciences (DCCPS) has previously demonstrated that approximately 60% of population based studies funded by the division use existing biospecimens from other collections, and that those studies are more cost and time efficient than studies collecting new specimens. Yet, it is difficult for researchers to identify potentially appropriate sources for biospecimens and accompanying epidemiologic and exposure data. Development of a searchable inventory of population-based biospecimen resources was a major recommendation resulting from an NCI think tank held in August 2013 (“Utilizing Existing Clinical and Population Biospecimen Resources for Discovery or Validation of Markers for Early Cancer Detection”) and would also be directly addressing four of the key recommendations that emerged in an NCI sponsored workshop titled “Trends in 21<sup>st</sup> Century Epidemiology: From Scientific Discoveries to Population Health” (*CEBP*, 2013, issue 22, page 508). In response to this, NCI DCCPS is developing a biospecimen inventory and online searchable catalog (or “Population Sciences Biospecimen Catalog (PSBC)”). The PSBC allows scientists in the research community and the NCI to locate specimens appropriate for their population based research projects. It is not NCI’s intent to collect biospecimens; rather the collections are descriptions of the available data that can act as a resource and be shared with researchers and scientists who are interested. This submission is via data upload to the secure website in order to collect information to manage and improve a program and its resources for the use by all scientists.

### **A.1 Circumstances Making the Collection of Information Necessary**

Section 410 of the Public Health Service Act (42 USC § 285) authorizes the collection of the information. The Division of Cancer Control and Population Sciences (DCCPS) and the National Cancer Institute (NCI) have been committed to fostering collaboration and sharing of resources. Over the past several years, it has become apparent that to promote such collaborations involving biospecimens for population research studies, investigators and program directors would greatly benefit from having a clearer understanding of

biospecimen-related epidemiologic resources funded by NCI. This is exemplified with specific recommendations from an August 2013 think tank (“Utilizing Existing Clinical and Population Biospecimen Resources for Discovery or Validation of Markers for Early Cancer Detection”) that was held by DCCPS and the Division of Cancer Prevention (DCP). The purpose of the meeting was to learn whether ongoing NCI-funded projects, e.g. cohort studies, HMOs, or clinical trials, might be leveraged for unbiased studies of biomarker study and validation. All participants at the think tank agreed that leveraging existing biospecimen resources from population based cohorts, RCTs, and case-control was essential not only for biomarker validation but also to extend the reach of epidemiology beyond etiology. However, during meeting discussions, it became apparent that participants were unaware of existing resources that were funded by NCI (and NIH) and what types of biospecimens were available for research purposes. One major recommendation resulting from the meeting was to create a searchable inventory of population-based biospecimen resources that were created using NCI funding. Such an inventory would also be directly addressing four of the key recommendations that emerged in an NCI sponsored workshop titled “Trends in 21<sup>st</sup> Century Epidemiology: From Scientific Discoveries to Population Health” (*CEBP*, 2013, issue 22, page 508).

Therefore, NCI’s Epidemiology and Genomics Research Program is creating a biospecimen inventory and online, searchable catalog (or “Population Sciences Biospecimen Catalog (PSBC)”). The PSBC allows scientists in the research community and the NCI to locate specimens appropriate for their population based research projects. It is not NCI’s intent to collect biospecimens; rather the collections are descriptions of the

available data that can act as a resource and be shared with researchers and scientists who are interested.

## **A.2 Purpose and Use of the Information Collection**

The purpose of this information collection is to characterize the biospecimen inventory of the respondents. The information collected will allow scientists to search the database for specimen collections with particular characteristics and accompanying data needed for their population based research study. We previously demonstrated that approximately 60% of population based studies funded by NCI DCCPS use existing biospecimens from biorepositories that were established for prior research studies, and that those study are more cost and time efficient than studies collecting new specimens. Yet, it is difficult for researchers to identify potentially appropriate sources for biospecimens and accompanying epidemiology, environmental, dietary (etc.) data.

The PSBC will provide a mechanism for investigators to easily search in a single location for such resources. However, in order to do so, researchers with such collections need to contribute the data to the PSBC. The respondents are principle investigators of NCI grants, biorepository managers, and study managers associated with the NCI grant. They will be sent an initial letter/email requesting they complete the information about the biospecimens in their inventory (**Attachment 2**). The letter includes a link to provide the requested information (**Attachment 1**), either as an upload (Excel file) or automatically through an Application Program Interface (API) data delivery to the PSBC.

The requested information includes:

- Specimen collection contact information
- Study design (e.g. case/control, cohort, survivor cohort, case only, family study, nested case-control)
- Collection features (e.g. metastases included)
- Specimen collection frequency (i.e. serial collection)
- Specimens (collected from which cancer site; specimen type, e.g. whole blood, serum, plasma, Guthrie cards/blood spots, etc; how many cases are specimens from; were specimens collected for this study)
- Additional data (e.g. dietary information collected, environmental exposures assessed, codebook of variable available, linkage to electronic medical records available, and more)

The information collected will be reviewed by the DCCPS Biospecimen Coordinator and the administrator of the website. The collected information will be used for the management of the website to benefit the research community. Additionally, the information will be available to investigators, via queries to the PSBC (via an existing EGRP web-based database, Cancer Genomics and Epidemiology Navigator (CGEN; <http://epi.grants.cancer.gov/cgen/>).

An annual update of the information will be requested by email of the respondents to ensure that the biospecimens inventory remains up to date, accurate and available **(Attachment 3)**. The respondents will be asked to log on to the PSBC and confirm that the information entered is still correct or, if needed, make changes via an excel file and resending it to CGEN. Additionally, the respondents have the option of simply rerunning their program code (which would have already been written for the initial request) to push their data through to the database electronically through an API. Majority of the

information will be pre-populated and thus a very short amount of time is needed to provide the update.

### **A.3 Use of Information Technology and Burden Reduction**

To reduce the burden for the respondents, several things will be done. First, the PSBC will be integrated with the existing CGEN web-based interface and database. Doing so will enable the PSBC to be pre-populated, when possible, with information pulled from NIH grants associated with the biospecimen collection. Second, several options will be given to respondents to allow them to provide the information in electronic formats. They may choose to provide the information in an Excel file format, or as a download from their own database. Information Management Services, Inc. (IMS), a NCI contractor, serves as the website administrator. They have developed an API to allow a specimen resource institute's IT personnel to create a code that can then be easily run to automatically provide updates to the PSBC.

The biospecimen catalog will be added as a tab on the existing CGEN website (<http://epi.grants.cancer.gov/cgen/>) that is hosted on the NCI EGRP webpage. The websites have already been assessed with a PIA and are approved.

### **A.4 Efforts to Identify Duplication and Use of Similar Information**

The PSBC is a unique website that collects detailed descriptions of population-based biospecimen resources and accompanying data, and makes them available to the scientific community. There are two existing NCI-resources that will be leveraged when



developing the PSBC: 1) Specimen Resource Locator (SRL) supported by the Division of Cancer Treatment and Diagnosis (DCTD), NCI and 2) the Cancer Genomics and Epidemiology Navigator (CGEN). The biospecimen catalog will be accessed through the CGEN interface and macro-level data from the biospecimen catalog will be added to the SRL. The PSBC captures more detailed information about the biospecimen collections compared with the SRL. Specifically, the PSBC includes the following information that is essential for population studies, but is not captured in the SRL: study design (e.g. case/control, cohort, survivor cohort, case only, family study, nested case-control), specimen collection frequency (i.e. serial collection), dietary information collected, environmental exposures assessed, codebook of variable available, linkage to electronic medical records available, physical activity assessed, genomic data available, lifestyle factors recorded, other non-cancer medical conditions recorded, clinical data available, and additional specimens (such as feces, Guthrie cards/blood spots, tissue culture, saliva for biomarkers); see **Attachment 1**.

The SRL does not accommodate the specific biospecimen and related data information associated with population-based studies. Therefore, NCI EGRP has developed the PSBC to capture the biospecimen and related data information that is essential for population-based researchers to search for appropriate biospecimen resources .

#### **A.5 Impact on Small Businesses or Other Small Entities**

No small businesses will be involved.

#### **A.6 Consequences of Collecting the Information Less Frequently**

The respondent will initially be asked to provide the information and then to update the information annually thereafter. The annual updates will keep the PSBC accurate and up to date on the inventory. The consequences of not updating the inventory may result in inaccurate inventory numbers, available specimens, and specimen access information. We will ease this burden by providing an API and/or an excel spreadsheet for our investigators.

#### **A.7 Special Circumstances Relating to the Guidelines of 5 CFR 1320.5**

This project will be implemented in a manner that fully complies with Guidelines of 5 CFR 1320.5

#### **A.8 Comments in Response to the Federal Register Notice and Efforts to Consult Outside Agency**

A 60-day Federal Register notice soliciting comments was published in the Federal Register on June 30, 2015, Vol. 80, P. 37280. No comments were received.

To create this biospecimen catalog, Jean Cyr (the NCI contractor for this project) at IMS (Information Management Services, Inc), the Project Manager for this project, was consulted; Jean's email is: [CyrJ@imsweb.com](mailto:CyrJ@imsweb.com).

#### **A.9 Explanation of Any Payment of Gift to Respondents**

No payment or gifts will be given to the respondents.

#### **A.10 Assurance of Confidentiality Provided to Respondents**

All information will be kept private to the extent allowable by law. Though name and contact information are collected, the biospecimen resource institutions are responding on behalf of their company not themselves and thus no personally identifiable information (PII) will be collected.

Since this is not considered research nor will there be publications, the Federal regulations for the protection of human subjects do not apply to this activity.

#### **A.11 Justification for Sensitive Questions**

No sensitive questions or PII are being collected.

#### **A.12 Estimates of Hour Burden Including Annualized Hourly Costs**

The annualized hour burden will be 80 hours to conduct both the initial request and an annual update (Table A.12-1) from 120 total participants. This amounts to approximately 240 hours over the three-year information collection phase. The respondents include private sector (business or other for-profits and not-for profits institutions), State and Federal Governments. A majority of the respondents are NCI funded grantees. The federal government respondents are responding as part of their work functions and are exempt under PRA. The annual update will be the prepopulated CGEN Population Sciences Biospecimen initial request and requires the principal investigators to only review and correct any outdated information.

Table A12-1. Estimated Annualized Burden Hours

Type of Respondent	Form Name	Number of Respondents	Number of Responses Per Respondent	Average Burden Per Response (in hours)	Total Burden Hour
Private Sector	Population Sciences	30	1	1	30
State Government	Biospecimen Catalog Initial Request	30	1	1	30
Private Sector	Population Sciences	30	1	20/60	10
State Government	Biospecimen Catalog Annual Update	30	1	20/60	10
Total		120			80

The annualized estimated cost to respondents is \$2,695.20; which is about \$8,085.60 over the three-year information collection phase (Table A.12-2). The hourly wage rate was obtained from the most recent data through the Bureau of Labor Statistics, [http://www.bls.gov/oes/current/oes\\_nat.htm#19-0000](http://www.bls.gov/oes/current/oes_nat.htm#19-0000), Occupation title “Life, Physical and Social Science Technicians”, occupation code 19-0000.

Table A12-2. Annualized Cost to Respondents

Type of Respondent	Number of Respondents	Total Annual Burden Hours	Wage Rate	Respondent Cost
Private Sector, State and Federal Governments	120	80	\$33.69	\$2,695.20

**A.13 Estimate of Other Total Annual Cost Burden to Respondents or Record Keepers**

There are no capital costs, operating costs, or maintenance costs to report.

**A.14 Annualized Cost to the Federal Government**

The annualized cost to the Federal Government is \$30,163.50. The Federal personnel cost is \$20,163.50. Federal personnel will oversee the deposition of information into the database. The contractor costs are \$10,000 ; (Table A.14-1).

Table A14-1 Annualized Cost to the Federal Government

<b>Staffing</b>	<b>Task</b>	<b>Annualized Cost</b>
NCI	Program Director, Grade 14, Step 5 (10% time for 12 months at \$ 121,635)	\$12,163.50
	Research Associate/Fellow (20% time for 12 months at \$40,000)	8,000
Contractor	Project Management Support, Web Management support	\$10,000
Total		\$30,163.50

**A.15 Explanation for Program Changes or Adjustments**

This is a new information collection.

**A.16 Plans for Tabulation and Publication and Project Time Schedule**

There are no plans for detailed statistical analysis of the information collected. However, descriptive analyses will be used to monitor the metrics such as total number of specimens by cancer site, environmental data, etc.; queries by cancer site, specimen type, and successful queries.

The project time schedule can be seen in Table A16-1.

**Table A16-1 Project Time Schedule**

Task	Months After OMB Approval
Web start up, design, content, URL	ongoing
Email potential respondents with the initial request	1-2 months for the first potential respondents; monthly afterwards for additional potential respondents
Review incoming electronic applications	2-3 months initially; then ongoing thereafter
Annual updates 2 times per year (March/Sept, depending on when the initial data was provided)	13-36 months

**A.17 Reason(s) Display of OMB Expiration Date is Inappropriate**

There is no request for exemption from displaying expiration date for OMB approval.

**A.18 Exceptions to Certification for Paperwork Reduction Act Submissions**

These data are collected in a manner consistent with the certification statement.

No exceptions are requested.