Data Collection Form
7/29/2014

OMB No.: 0925-XXXX
Expiration Date: xx/xx/20xx

OMB No.: 0925-XXXX

Expiration Date: xx/xx/20xx

Collection of this information is authorized by The Public Health Service Act, Section 411 (42 USC 285a). Rights of study participants are protected by The Privacy Act of 1974. Participation is voluntary, and there are no penalties for not participating or withdrawing from the study at any time. Refusal to participate will not affect your benefits in any way. The information collected in this study will be kept private to the extent provided by law. Names and other identifiers will not appear in any report of the study. Information provided will be combined for all study participants and reported as summaries. You are being contacted by email to complete this instrument so that we can develop a database and keep the website up to date.

Public reporting burden for this collection of information is estimated to average 45 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. **An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number.** Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to: NIH, Project Clearance Branch, 6705 Rockledge Drive, MSC 7974, Bethesda, MD 20892-7974, ATTN: PRA (0925-XXXX). Do not return the completed form to this address.

Thank you for taking the time to complete this form. The information you provide will populate the Cancer Epidemiology Descriptive Cohort Database (give URL). Users of the CEDCD will be able to find information about Cancer Epidemiology Cohorts such as yours in a single unified database. The CEDCD will enable users to learn about existing cohorts, compare cohort characteristics, and tabulate counts of participants, cancers, and specimens across cohorts. We hope you will find the CEDCD useful in seeking collaborators and facilitating projects.

This form is pre-filled with as much information as was possible to locate from available sources. Please review for accuracy and add information as needed.

Please return this form to Westat (cedcdhelpdesk@westat.com). The information on this form will be electronically loaded to the CEDCD database through an automated process. Annual updates are planned to ensure that the database reflects accurate up-to-date information about your cohort.

Data Collection Form
7/29/2014

OMB No.: 0925-XXXX
Expiration Date: xx/xx/20xx

A. Basic Cohort Information *(If your cohort is comprised of more than one distinct enrollment period (such as Physicians Health Study I and II), please complete separate Cohorts Descriptive Database Collection Forms to treat them as separate cohorts.)*

A.1a Cohort Name:	
A.1b Cohort Abbreviation:	
A.1c Cohort Website: (if available)	
A.2 Date Form Completed:	_____ MM / DD / YYYY

A.3 Person whom completed the form:

Name:	_____
Position on the cohort:	_____
Phone:	_____
Email:	_____
Is this the person to contact with questions about this form?	☐ No ☐ Yes If no, please provide the name and contact information for correct person in the space on the right.

Contact Person for clarification of this form:

Name:	_____
Position on the cohort:	_____
Phone:	_____
Email:	_____

A.4 Cohort's Principal Investigator(s) and Co-Investigators:
If there is not enough room below to list all of the investigators, please attach a separate document listing all of the investigators with all the information specified below. Please provide title at your home institution.

Name:	_____	Name:	_____
Title:	_____	Title:	_____
Institution:	_____	Institution:	_____
Phone:	_____	Phone:	_____
Email:	_____	Email:	_____
Name:	_____	Name:	_____
Title:	_____	Title:	_____
Institution:	_____	Institution:	_____
Phone:	_____	Phone:	_____
Email:	_____	Email:	_____

	_____		_____

A. Basic Cohort Information (continued)

A.5 If an investigator is interested in collaborating with your cohort on a new project, whom should they contact?

Name:	_____
Position on the cohort:	_____
Phone:	_____
Email:	_____

A.6 Cohort Description:

Please provide a short paragraph describing your cohort. This will be used as an overall narrative description of your cohort on your cohort's page on the CEDCD website.

A. Basic Cohort Information (continued)

A.7 Cohort Design:		Check one: ☐ Risk Cohort – *(initially enrolled participants without cancer)* ☐ Survivor Cohort – *(initially enrolled participants with cancer)* ☐ Lifecycle Cohort – (*multi-generational enrollment within families*)
A.8 Is the cohort a survivor cohort built from a previously established risk cohort?	☐ No ☐ Yes	If yes, Were data collected before enrollment into the survivor cohort? ☐ No ☐ Yes Were biospecimens collected before enrollment into the survivor cohort? ☐ No ☐ Yes Please complete the remainder of this form as it pertains only to data and specimens collected from establishment of the survivor cohort; do not include data and specimens collected as part of the previously established cohort.

A.9 Is this a multi-site cohort? ☐ No ☐ Yes

If yes, please list the participating institutions and investigators:

a. Investigator: _____ Institution: _____
b. Investigator: _____ Institution: _____
c. Investigator: _____ Institution: _____
d. Investigator: _____ Institution: _____
e. Investigator: _____ Institution: _____
f. Investigator: _____ Institution: _____
g. Investigator: _____ Institution: _____
h. Investigator: _____ Institution: _____
i. Investigator: _____ Institution: _____
j. Investigator: _____ Institution: _____

A. Basic Cohort Information (continued)	
A.10 Eligibility Criteria:	Age: Eligible Age Range: _____ to _____ Gender: ☐ Both genders eligible ☐ Males only eligible ☐ Females only eligible
A.11 Enrollment:	_____ to _____ Year Started (YYYY) Year Ended (YYYY) ☐ Check if enrollment is ongoing
A.12 Age at Enrollment (range and median):	Range: _____ To _____ Median: _____ If your cohort is a lifecycle cohort enrolling multiple generations within families, then specify the age of each generation. First Generation – Age Range: _____ To _____ Second Generation – Age Range: _____ To _____ Third Generation – Age Range: _____ To _____
A.13 Specify time intervals when your questionnaire data were collected. For example, yearly, biannually, 2011-2013.	Specify:
A.14 Most recent year when questionnaire data were collected:	_____ Year (YYYY)
A.15 How was information from the questionnaire administered/collected?	Check all that apply: ☐ In person ☐ Paper ☐ Electronic / Web-based ☐ Other, specify: _____ _____
A.16 Were any tools aside from questionnaires used for exposure data collection? (e.g., an accelerometer for recording physical activity)	☐ No ☐ Yes If yes, specify the instruments: _____ _____
A.17 Most recent year of confirmed cancer case ascertainment:	_____ Year (YYYY)

A. Basic Cohort Information (continued)	
A.18 Most recent year of mortality follow-up:	_____ Year (YYYY)
A.19 Does your cohort have any known restrictions on participating in collaborative projects involving pooling of data or specimens or use of specimens in genomic studies? (For example, restrictions due to the wording of the informed consent?)	☐ No ☐ Yes If Yes, please describe briefly: _____ _____

B. Current Enrollment Counts

B.1 Total number of subjects enrolled: _____

If still enrolling, please specify the target number you plan to enroll: _____

If still enrolling, please specify by when do you plan to enroll subjects: _____ YYYY

B.2 Number of Males enrolled: _____

B.3 Number of Females enrolled: _____

B. 4 Racial Categories	Ethnic Categories									**Total**
	Not Hispanic or Latino			Hispanic or Latino			Unknown/Not Reported Ethnicity			
	Female	Male	Unknown/ Not Reported	Female	Male	Unknown/ Not Reported	Female	Male	Unknown/ Not Reported	
American Indian/ Alaska Native										
Asian										
Native Hawaiian or Other Pacific Islander										
Black or African American										
White										
More Than One Race										
Unknown or Not Reported										
Total										

C. Data on Major Content Domains

Specify whether you collected data within these major content domains. Baseline refers to data collected at or near enrollment into the cohort. If a lifecycle cohort, include all exposure data for all generations as follow-up.

Did you collect data on:	Check all that apply	
	Collected at baseline	**Collected during follow-up**
C.1 Marital Status	☐ Collected	☐ Collected
C.2 Socio-economic status (e.g., income)	☐ Collected	☐ Collected
C.3 Education Level	☐ Collected	☐ Collected
C.4 Anthropometry (e.g., weight, height, waist circumference, or BMI)	☐ Collected	☐ Collected
C.5 Cigarette smoking	☐ Collected	☐ Collected
C.6 Use of tobacco products other than cigarettes	☐ Collected If collected, specify other tobacco products: ☐ Cigars ☐ Pipes ☐ Chewing tobacco ☐ Other, specify: _____ _____	☐ Collected If collected, specify other tobacco products: ☐ Cigars ☐ Pipes ☐ Chewing tobacco ☐ Other, specify: _____ _____
C.7 Alcohol consumption	☐ Collected	☐ Collected
C.8 Dietary intake	☐ Collected	☐ Collected
C.9 Dietary supplement use	☐ Collected	☐ Collected
C.10 Physical activity	☐ Collected	☐ Collected
C.11 Reproductive history	☐ Collected	☐ Collected

C.12 Quality of life or other psychosocial variables	☐ Collected	☐ Collected

C. Data on Major Content Domains (continued)

Did you collect data on:	Check all that apply: Collected at baseline	Check all that apply: Collected during follow-up
C.13 Prescription medication use (not related to cancer treatment)	☐ Collected	☐ Collected
C.14 Non-prescription medication use (not related to cancer treatment)	☐ Collected	☐ Collected
C.15 Family history of cancer	☐ Collected If collected, was data collected on: ☐ First degree relatives only ☐ First and second degree relatives ☐ All relatives Do you have pedigrees? ☐ No ☐ Yes	☐ Collected If collected, was data collected on: ☐ First degree relatives only ☐ First and second degree relatives ☐ All relatives Do you have pedigrees? ☐ No ☐ Yes
C.16 Environmental or occupational exposures (e.g. air contaminants/quality, occupational exposures and history, water source)	☐ Collected	☐ Collected
C.17 Geocoding Information	☐ Collected	☐ Collected

C.18 Non-Cancer Illness: Check all that apply.

Illness	Prevalent Illness	Incident Illness
a. Diabetes	☐ Collected	☐ Collected
b. Heart and Vascular Diseases	☐ Collected	☐ Collected
c. Lung Diseases	☐ Collected	☐ Collected
d. Digestive and/or Genitourinary	☐ Collected	☐ Collected

Diseases		
e. Osteoporosis/Bone related conditions	☐ Collected	☐ Collected

C. Data on Major Content Domains (continued)

C.18 Non-Cancer Illness: Check all that apply.

Illness	Prevalent Illness	Incident Illness
f. Neurodegenerative Disorders and/or Mental Illnesses	☐ Collected	☐ Collected
g. Autoimmune diseases	☐ Collected	☐ Collected

D. Cancer Information: This section is to capture the extent of cancer information that your cohort collects or currently has available. Please limit your response to data that have already been collected or are part of ongoing collection, and not to include planned collection that has yet to begin. Please include in your consideration data that you ascertain from participants or other data sources as well as derived data (e.g. algorithms to differentiate recurrent vs second primary cancer).

D.1 How were your cancer cases ascertained?		Check all that apply: ☐ Self-report ☐ Tumor registry ☐ Medical record review ☐ Other, specify: _____ _____
D.2 Do you have recurrent cancer diagnosis?		☐ No ☐ Yes
D.3 Do you have second primary cancer diagnosis?		☐ No ☐ Yes
D.4 Do you have cancer treatment data?	☐ No ☐ Yes If no, would it be possible to collect this information from medical records or other sources? ☐ No ☐ Yes	If yes, specify treatment and data source: Treatment: Check all that apply ☐ Surgery ☐ Radiation ☐ Chemotherapy ☐ Hormonal therapy ☐ Bone marrow/stem cell transplant ☐ Other, specify: _____ _____ Data source: Check all that apply

		☐ Administrative claims data ☐ Electronic record ☐ Chart abstraction ☐ Patient-reported questionnaire ☐ Other, specify: _____

D. Cancer Information (continued)

D.5 Do you have cancer staging data?	☐ No ☐ Yes
D.6 Do you have tumor grade data?	☐ No ☐ Yes
D.7 Do you have tumor genetic markers data?	☐ No ☐ Yes If yes, please describe: _____ _____
D.8 Were cancer cases histologically confirmed?	Select only one: ☐ All ☐ Some ☐ None
D.9 Do you have cancer subtyping?	Check all that apply: ☐ Histological ☐ Molecular
D.10 Do you have information on cancer-related conditions?	☐ No ☐ Yes

If yes, check all that apply:

☐ Acute treatment-related toxicity (e.g., diarrhea, nephrotoxicity)

☐ Late effects of treatment (e.g., cardiotoxicity, lymphedema)

☐ Symptoms management (e.g., fatigue, pain, sexual dysfunction)

☐ Other, specify:

_____
_____

	☐ No ☐ Yes

D.11 If you **did not collect** the information requested in D.2 to D.10, are the data available to be retrieved at a later point in time?	

E. Mortality		
E.1 How was death confirmed by your cohort?		Check all that apply: ☐ National Death Index (NDI) linkage ☐ State death certificates ☐ Other, specify _____ _____
E.2 Do you have date of death for most subjects?		☐ No ☐ Yes
E.3 Do you have cause of death for most subjects?	☐ No ☐ Yes	If yes, is the cause of death coded? ☐ No ☐ Yes If yes, what type of death code was used? Check all that apply: ☐ ICD-9 ☐ ICD-10 ☐ Other, specify _____ _____
E.4 What is the number of deaths in your cohort as of most recent mortality follow-up?		_____

F. Data Linkage and Harmonization		
F.1 Have you linked your cohort data to any other existing databases (e.g., Center for Medicare and Medicaid Services, Surveillance, Epidemiology and End Results)?	☐ No ☐ Yes	If yes, specify:
F.2 Has your cohort participated in any cross-cohort data harmonization projects not limited to NCI?	☐ No ☐ Yes	If yes, specify:

G. Specimens Collected		
Specify the types of specimens you collected, whether the specimen was collected at baseline, and/or collected at other time points.		
Did you collect any of the following specimens:	**Collected at baseline**	**Collected at other time points**
G.1 Blood	☐ Collected If collected, types of aliquots (Check all that apply): ☐ Serum ☐ Buffy Coat ☐ Plasma ☐ Other Blood Derivative	☐ Collected If collected, types of aliquots (Check all that apply): ☐ Serum ☐ Buffy Coat ☐ Plasma ☐ Other Blood Derivative
G.2 Buccal Swab	☐ Collected	☐ Collected
G.3 Saliva	☐ Collected	☐ Collected
G.4 Lymphocytes	☐ Collected	☐ Collected
G.5 Other Specimen types not listed above (e.g., urine, sputum). Do not include tumor tissue. Specify below:		
a.	☐ Collected	☐ Collected
b.	☐ Collected	☐ Collected
c.	☐ Collected	☐ Collected
G.6 Did you collect tumor tissue?	☐ No (Skip to question G.10) ☐ Yes	
G.7 Did you also collect normal tissue?	☐ No ☐ Yes	
G.8 How were the tumor tissue samples prepared/ stored?	Check all that apply: ☐ Formalin Fixed Paraffin Embedded (FFPE) ☐ Fresh/Flash Frozen ☐ Diagnostic slides ☐ Other, specify:	

	_____ _____

G. Specimens Collected (continued)

G.9 How was the tumor tissue collected? (Include collection of the same tumor at different time points)	Check all that apply: ☐ Core Biopsy ☐ Fine Needle Aspirations (FNA) ☐ Surgery ☐ Other, specify: _____ _____
G.10 If your cohort does not currently collect tumor blocks, did you collect information on where the blocks are kept/stored?	☐ No ☐ Yes

Do you have:		Approximately How Many Participants
G.11 Genotyping Data (SNP)	☐ No ☐ Yes	
G.12 Sequencing Data – Exome	☐ No ☐ Yes	
G.13 Sequencing Data – Whole Genome	☐ No ☐ Yes	
G.14 Epigenetic or metabolic markers	☐ No ☐ Yes	
G.15 Other “omics” data	☐ No ☐ Yes	

H. Technology Use

H.1 In your cohort, have you adopted the use of mobile devices (i.e., tablet computers, personal digital assistants, etc.) for the collection and/or measurement of demographic or lifestyle factors, environmental exposures, and/or other types of information?	☐ Yes, please list or describe: _____ _____ _____ ☐ No, but we are currently considering it or will consider it in our next renewal. ☐ No, and we do not have any immediate plans to do so.
H.2 Most studies store all of their study data on local servers that are maintained at their institution. Cloud computing refers to storing data on the internet. Have you adopted the use	☐ Yes, please list or describe: _____ _____ _____

of cloud-based approaches for the collection, management, or distribution of any of your study data?	☐ No, but we are currently considering it or will consider it in our next renewal. ☐ No, and we do not have any immediate plans to do so.

H. Technology Use (continued)

H.3 If the answers were "No, and we do not have immediate plans to do so" for either of the prior 2 questions, please indicate the possible reasons.	Check all that apply: ☐ Limited funding ☐ Limited support from department/institution ☐ Limited technical infrastructure or support ☐ Security concerns ☐ Other and/or please describe: _____ _____ _____

I. Additional Items for Inclusion on the CEDCD Website

As indicated on the CEDCD Approval Form, we are requesting the following items for inclusion on the CEDCD website. If you provided approval to post this information, please attach the documents and return them to Westat with this form If they are already available on a publicly accessible website, please just provide the website address.

Document	Attached	Website URL (if document is not attached)
Questionnaires	☐	URL:
Main cohort protocol	☐	URL:
Data sharing policy	☐	URL:
Biospecimen sharing policy	☐	URL:
Publication (authorship) policy	☐	URL:
CEDCD Biospecimen and Cancer Count Information Spreadsheet	☐	Attached Only
Separate List of investigators (if needed)	☐	Attached Only