Supplementary Instrument 1. The Data Collection Form completed by cohort investigators

CEDCD Cancer Epidemiology Descriptive Cohort Database

Data Collection Form
02/20/2015

OMB No.: 0925-0706
Expiration Date: 11/30/2017

OMB No.: 0925-0706

Expiration Date: 11/30/2017

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 90 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 (0925-0706). 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 (http://CEDCD.nci.nih.gov). 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
02/20/2015

OMB No.: 0925-0706
Expiration Date: 11/30/2017

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

Please list the recruitment catchment areas (The catchment areas are defined as the geographical location from where participants are recruited):

a. Catchment area: _____
b. Catchment area: _____
c. Catchment area: _____
d. Catchment area: _____
e. Catchment area: _____
f. Catchment area: _____
g. Catchment area: _____
h. Catchment area: _____
i. Catchment area: _____
j. Catchment area: _____

A.10 Eligibility Criteria:	Age: Eligible Age Range: _____ to _____ Gender: ☐ Both genders eligible ☐ Males only eligible ☐ Females only eligible

A. Basic Cohort Information (continued)	
A.11 Enrollment:	_____ to _____ Year Started (YYYY) Year Ended (YYYY) Is enrollment ongoing? ☐ No ☐ Yes
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?	In person ☐ No ☐ Yes Paper ☐ No ☐ Yes Electronic/Web-based ☐ No ☐ Yes Other ☐ No ☐ Yes, 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.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: _____
Year (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:	Collected at baseline	Collected during follow-up
C.1 Marital Status	☐ No ☐ Yes	☐ No ☐ Yes
C.2 Socio-economic status (e.g., income)	☐ No ☐ Yes	☐ No ☐ Yes
C.3 Education Level	☐ No ☐ Yes	☐ No ☐ Yes
C.4 Anthropometry (e.g., weight, height, waist circumference, or BMI)	☐ No ☐ Yes	☐ No ☐ Yes
C.5 Cigarette smoking	☐ No ☐ Yes	☐ No ☐ Yes
C.6 Use of tobacco products other than cigarettes	☐ No ☐ Yes If collected, specify other tobacco products: Cigars ☐ No ☐ Yes Pipes ☐ No ☐ Yes Chewing tobacco ☐ No ☐ Yes Other ☐ No ☐ Yes, specify: _____	☐ No ☐ Yes If collected, specify other tobacco products: Cigars ☐ No ☐ Yes Pipes ☐ No ☐ Yes Chewing tobacco ☐ No ☐ Yes Other ☐ No ☐ Yes, specify: _____
C.7 Alcohol consumption	☐ No ☐ Yes	☐ No ☐ Yes
C.8 Dietary intake	☐ No ☐ Yes	☐ No ☐ Yes
C.9 Dietary supplement use	☐ No ☐ Yes	☐ No ☐ Yes
C.10 Physical activity	☐ No ☐ Yes	☐ No ☐ Yes
C.11 Reproductive history	☐ No ☐ Yes	☐ No ☐ Yes
C.12 Quality of life or other psychosocial variables	☐ No ☐ Yes	☐ No ☐ Yes
C.13 Prescription medication use (not related to cancer treatment)	☐ No ☐ Yes	☐ No ☐ Yes

C. Data on Major Content Domains (continued)		
Did you collect data on:	**Collected at baseline**	**Collected during follow-up**
C.14 Non-prescription medication use (not related to cancer treatment)	☐ No ☐ Yes	☐ No ☐ Yes
C.15 Family history of cancer	☐ No ☐ Yes If collected, were data collected on: First degree relatives only ☐ No ☐ Yes First and second degree relatives ☐ No ☐ Yes All relatives ☐ No ☐ Yes Do you have pedigrees? ☐ No ☐ Yes	☐ No ☐ Yes If collected, were data collected on: First degree relatives only ☐ No ☐ Yes First and second degree relatives ☐ No ☐ Yes All relatives ☐ No ☐ Yes Do you have pedigrees? ☐ No ☐ Yes
C.16 Environmental or occupational exposures (e.g., air contaminants/quality, occupational exposures and history, water source)	☐ No ☐ Yes	☐ No ☐ Yes
C.17 Geocoding Information	☐ No ☐ Yes	☐ No ☐ Yes
C.18 Non-Cancer Medical Conditions:		
Did you collect data on:	**Prevalent Medical Condition**	**Incident Medical Condition**
a. Diabetes	☐ No ☐ Yes	☐ No ☐ Yes
b. Heart and Vascular Diseases	☐ No ☐ Yes	☐ No ☐ Yes
c. Lung Diseases	☐ No ☐ Yes	☐ No ☐ Yes
d. Digestive and/or Genitourinary Diseases	☐ No ☐ Yes	☐ No ☐ Yes
e. Osteoporosis/Bone related conditions	☐ No ☐ Yes	☐ No ☐ Yes
f. Neurodegenerative Disorders and/or Mental Illnesses	☐ No ☐ Yes	☐ No ☐ Yes
g. Autoimmune diseases	☐ No ☐ Yes	☐ No ☐ Yes

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?		Self-report ☐ No ☐ Yes Tumor registry ☐ No ☐ Yes Medical record review ☐ No ☐ Yes Other ☐ No ☐ Yes, 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: Surgery ☐ No ☐ Yes Radiation ☐ No ☐ Yes Chemotherapy ☐ No ☐ Yes Hormonal therapy ☐ No ☐ Yes Bone marrow/stem cell transplant ☐ No ☐ Yes Other ☐ No ☐ Yes, specify: _____ Data source: Administrative claims data ☐ No ☐ Yes Electronic record ☐ No ☐ Yes Chart abstraction ☐ No ☐ Yes Patient-reported questionnaire ☐ No ☐ Yes Other ☐ No ☐ Yes, 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?	Histological ☐ No ☐ Yes Molecular ☐ No ☐ Yes
D.10 Do you have information on cancer-related conditions?	☐ No ☐ Yes
If yes, specify the information on the cancer related conditions below: Acute treatment-related toxicity (e.g., diarrhea, nephrotoxicity) ☐ No ☐ Yes Late effects of treatment (e.g., cardiotoxicity, lymphedema) ☐ No ☐ Yes Symptoms management (e.g., fatigue, pain, sexual dysfunction) ☐ No ☐ Yes Other ☐ No ☐ Yes, specify _____ _____	
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?	☐ No ☐ Yes

E. Mortality		
E.1 How was death confirmed by your cohort?		National Death Index (NDI) linkage ☐ No ☐ Yes State death certificates ☐ No ☐ Yes Other ☐ No ☐ Yes, 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? ICD-9 ☐ No ☐ Yes ICD-10 ☐ No ☐ Yes Other ☐ No ☐ Yes, 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	☐ No ☐ Yes If collected, types of aliquots Serum ☐ No ☐ Yes Buffy Coat ☐ No ☐ Yes Plasma ☐ No ☐ Yes Other Blood Derivative ☐ No ☐ Yes	☐ No ☐ Yes If collected, types of aliquots Serum ☐ No ☐ Yes Buffy Coat ☐ No ☐ Yes Plasma ☐ No ☐ Yes Other Blood Derivative ☐ No ☐ Yes
G.2 Buccal Swab	☐ No ☐ Yes	☐ No ☐ Yes
G.3 Saliva	☐ No ☐ Yes	☐ No ☐ Yes
G.4 Lymphocytes	☐ No ☐ Yes	☐ No ☐ Yes
G.5 Other Specimen types not listed above (e.g., urine, sputum). Do not include tumor tissue. Specify below:		
a.	☐ No ☐ Yes	☐ No ☐ Yes
b.	☐ No ☐ Yes	☐ No ☐ Yes
c.	☐ No ☐ Yes	☐ No ☐ Yes

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. Specimens Collected (continued)	
G.8 How were the tumor tissue samples prepared/ stored?	Formalin Fixed Paraffin Embedded (FFPE) ☐ No ☐ Yes Fresh/Flash Frozen ☐ No ☐ Yes Diagnostic Slides ☐ No ☐ Yes Other ☐ No ☐ Yes, specify: _____ _____
G.9 How was the tumor tissue collected? (Include collection of the same tumor at different time points)	Core Biopsy ☐ No ☐ Yes Fine Needle Aspirations (FNA) ☐ No ☐ Yes Surgery ☐ No ☐ Yes Other ☐ No ☐ Yes, 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:		**Specify approximately how many participants, case/control, and the cancer type for each data type below:** (attach separate sheet if you need more space)
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 of cloud-based approaches for the collection, management, or distribution of any of your study data?	☐ 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.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.	Limited funding ☐ No ☐ Yes Limited support from department/institution ☐ No ☐ Yes Limited technical infrastructure or support ☐ No ☐ Yes Security concerns ☐ No ☐ Yes Other ☐ No ☐ Yes, 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

Supplementary Instrument 2. The Biospecimen and Cancer Count Information Spreadsheet completed by cohort investigators

CEDCD Cancer Epidemiology Descriptive Cohort Database

Biospecimen and Cancer Count Information Spreadsheet

OMB No.: 0925-0706
Expiration Date: 11/30/2017

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 60 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 (0925-0706). Do not return the completed form to this address.

Instructions: Please complete the table on the next page as specified below:

In row 1, provide your cohort name.

In row 2, specify if the Cancers are incident or prevalent cancers. If you want to provide us with both incident and prevalent cancers, please complete two Biospecimen and Cancer Count Information Spreadsheet and specify the cancer type at the top. For this study, incident cancers are defined as cancer diagnosed after enrollment for all cohort types (risk, survival or life-cycle generations). Prevalent cancers are defined as cancers prior to enrollment in the study. It would include the cancer that was used as an eligiblity criteria into the survival cohort.

In row 3, please provide any notes that would help clarify the information you are providing in the 'Notes' section.

In Row 4, enter the number of males and females, and blood, tumor tissue and other specimens in your current biospecimen inventory that do not have cancer. Tumor tissue FFPE is Formalin Fixed Paraffin Embedded.

In Rows 5-28, enter the number of males and females, and blood, buccal, feces, lymphocytes, tumor tissue, sputum, urine, and other specimens in your current biospecimen inventory by the cancer type listed .

In 'Other specimens' please enter the name and count of all specimens other than the ones specified.

If you do not have exact counts, please enter approximate counts.

If it is easier, you can send us this information in the format below in another file type such as .txt, .csv, .sas, or another similar data file. Contact us if you have questions about the type of other data files we will accept.

Questions? Please contact CEDCD Helpdesk at Westat (CEDCDHelpdesk@westat.com or 240-314-5860).

CEDCD Cancer Epidemiology Descriptive Cohort Database

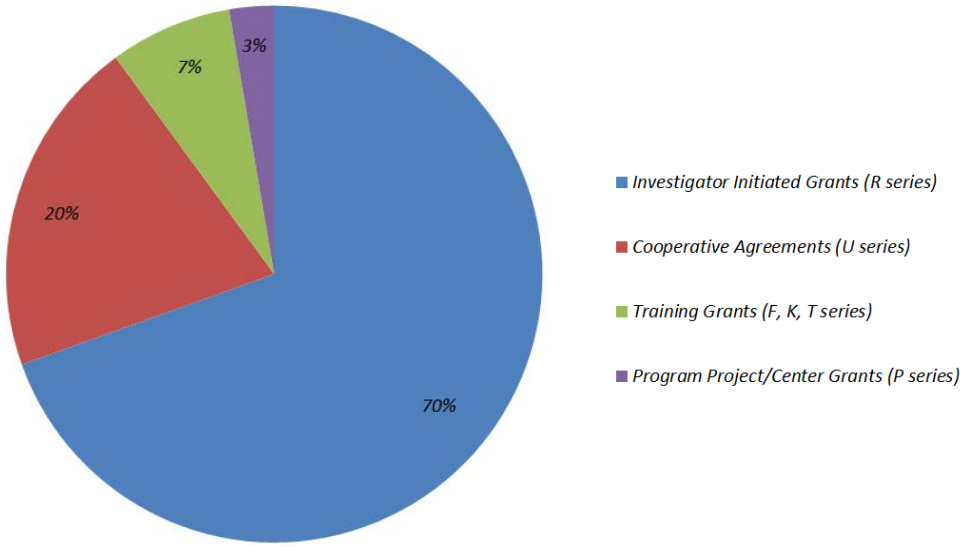
Biospecimen and Cancer Count Information Spreadsheet

1	**Cohort Name: Breast Cancer Detection Demonstration Project Follow-up Study (BCDDP)**														
2	**Specify if the cancers are Incident Cancers or Prevalent Cancers: Incident**														
3	**Notes:**														
	ICD-9	**ICD-10/O**	**Cancer Type**	**Males**	**Females**	**Blood**	**Buccal**	**Feces**	**Lymphocytes**	**Tumor Tissue Fresh/Frozen**	**Tumor Tissue FFPE**	**Sputum**	**Urine**	**DNA**	**Other Specify:**
4			No Cancer												
5	141-149	C00-C14	Oropharyngeal			0								0	
6	150	C15	Esophagus			0								0	
7	151	C16	Stomach			0								0	
8	152	C17	Small intestine												
9	153	C18	Colon			0								0	
10	154	C19-C21	Rectum and anus			0								0	
11	155	C22	Liver and intrahepatic bile ducts			0								0	
12	156	C23, C24	Gall bladder and extrahepatic bile duct			0								0	
13	157	C25	Pancreas			0								0	
14	162	C33, C34	Trachea, bronchus, and lung			0								0	
15	170	C40	Bone			0								0	
16	172	C43	Melanoma (excluding genital organs)			0								0	
17	174-175	C50	Breast			0								0	
18	180	C53	Cervix			0								0	
19	182	C54	Corpus, body of uterus			0								0	
20	183	C56	Ovary, fallopian tube, broad ligament			0								0	
21	185	C61	Prostate			0								0	
22	188	C67	Bladder			0								0	
23	189	C64-C66, C68	Kidney and other unspecified urinary organs including renal pelvis, ureter, urethra			0								0	
24	191	C71	Brain			0								0	
25	193	C73	Thyroid			0								0	
26	200-202	C81-C85	Lymphoma (HL and NHL)			0								0	
27	203	C90	Myeloma			0								0	
28	204-208	C91-C95	Leukemia			0								0	

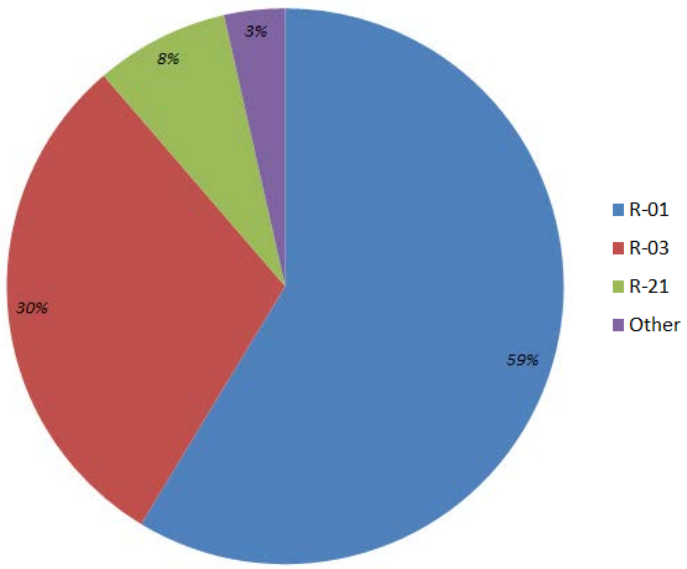
Supplementary Table 1. Number of cohorts with participants from the catchment area. The catchment areas are defined as the geographical location from where participants are recruited.

Country	Number of Cohorts
United States	30
Canada	8
Australia	3
China	3
Norway	2
Sweden	2
United Kingdom	2
Denmark	1
Finland	1
France	1
Greece	1
Iran	1
Italy	1
Japan	1
Puerto Rico	1
Singapore	1
Spain	1
The Netherlands	1

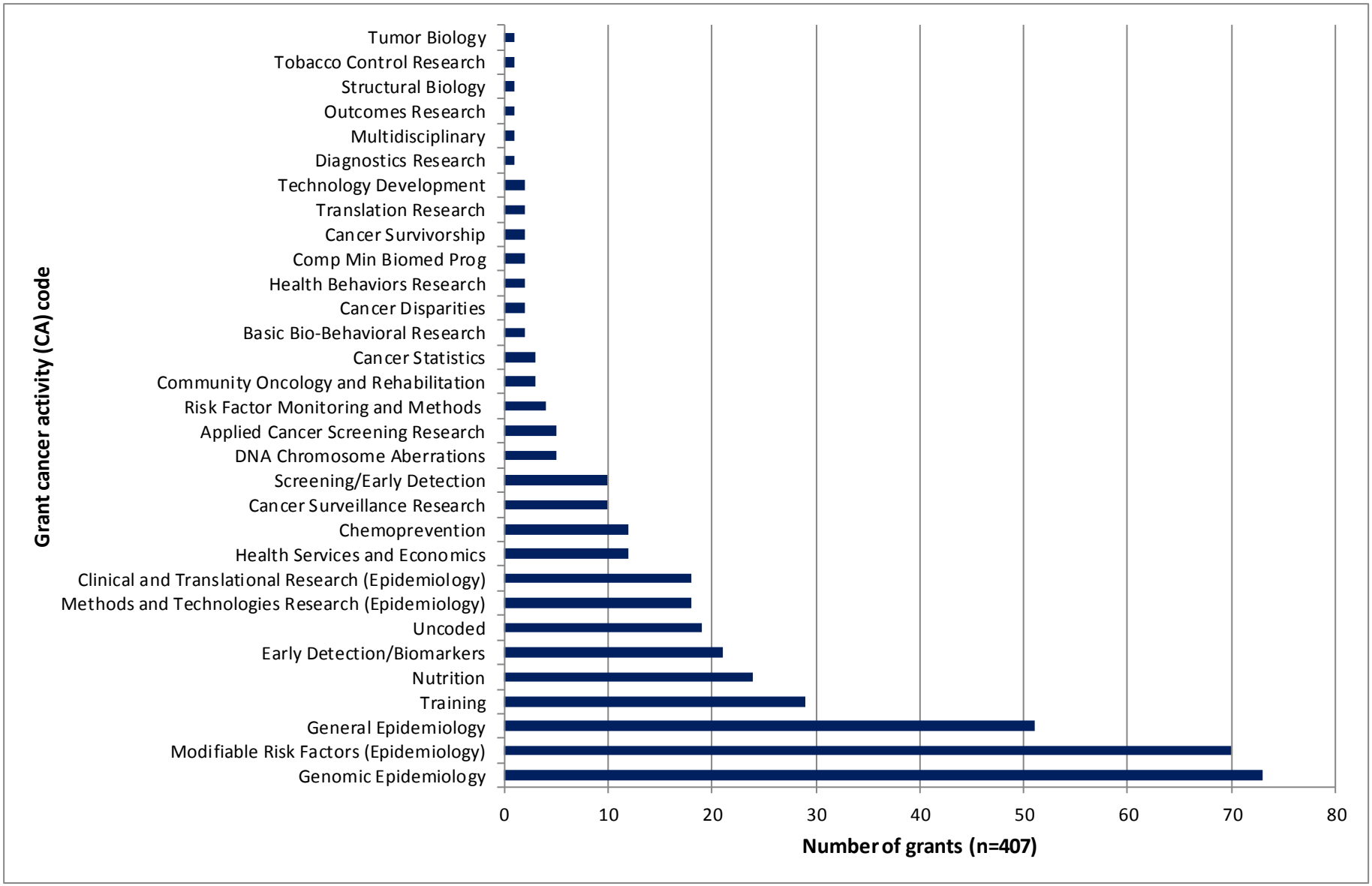
A)



B)



Supplementary Figure 1. Breakdown of NCI-awarded grants to participating cohorts. **A)** Funding mechanism of awarded grants, shown as percent of total NCI awarded grants (n=407). **B)** Categorization of investigator initiated grants funded by NCI, shown as percent of R-series awarded grants (n=283).



Supplementary Figure 2. Distribution of NCI-funded grants (n=407) categorized by cancer activity code. Data retrieved from Portfolio Management Application (PMA) 16.1 on Wednesday August 19, 2015.

Supplementary Figure 3. Screenshot of the home page of the CEDCD (https://cedcd.nci.nih.gov/)