**APPLICATION FOR SEER-MEDICARE LINKED DATA - 2022 LINKAGE**

(Please complete all sections and provide all required information)

|  |  |
| --- | --- |
| PROJECT TITLE\* |  |

\*This title must match the project title listed in the IRB letter and signed DUA

**SECTION I: Contact information**

Principal Investigator: (*students or fellows may NOT be listed as the PI)*

|  |  |
| --- | --- |
| Name: |  |
| Institution: |  |
| Address: |  |
| City, State Zip |  |
| Email: |  |
| Phone |  |

Alternate contact: (Identify role: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)

|  |  |
| --- | --- |
| Name: |  |
| Institution: |  |
| Address: |  |
| City, State Zip |  |
| Email: |  |
| Phone: |  |

**SECTION II. Project Description:**

A. BRIEF OVERVIEW (one or two sentences):

B CANCER(S) OF INTEREST (refer to [Site Recode ICD-O-3/WHO 2008 List](https://seer.cancer.gov/siterecode/icdo3_dwhoheme/)):

C. DESCRIPTION OF PROJECT (between 1-5 pages). This description **must** include:

* statement of main hypothesis/research question
* brief scientific background for the study and the need to evaluate this population
* description of study subjects and cancer sites/phases to be included in the analysis
* Consider that beneficiaries “[likely to have complete claims](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7225666/)” are typically 65+ and

 continuously enrolled in part A/B with no HMO enrollment during the study period

* brief explanation of how key components of the study will be obtained/identified within the cancer and/or claims data– specifically:
* cohort selection criteria
* covariates
* outcomes
* Some variables have limitations, a list of which can be found [here.](https://healthcaredelivery.cancer.gov/seermedicare/considerations/measures.html) If you still intend to use these variables, please speak to how you will acknowledge or address their challenges.
* a table summarizing how each requested file will be used (e.g., PDE will be used…NCH will be used…)
* details about how 5% populations (non-cancer and/or cancer) will be used, if requested
* if you will link to any external datasets, provide details on how and why you will do so, including the variables of interest in the linked data and why the linkage is necessary (e.g., the data are not available in the current SEER-Medicare linkage)
* description of the planned analyses for each study aim (e.g., cox regression will assess…)
* description of personnel involved, and an explicit list of who will have access to the data
* timeline for completion, typically at least one year due to the complexity of the dataset (e.g., Month 1-3: obtain data. Month 3-4: Create analytic file…)
* brief description of the research output and how it will be disseminated/made available
* references can be included, if relevant

D. DATA STORAGE AND PROTECTION:

The **preferred method** of data storage is on an institutional server with all the protections that provides. It is strongly recommended that a plan showing data storage on an institutional server is provided.

If cloud-based storage is proposed, approval of a DMP-SAQ is required:

1. If the investigator has previously requested data from CMS and received CMS approval of a DMP for a cloud-based storage environment, then it can be re-used for storing SEER-Medicare data. In this case, a letter confirming CMS approval of the DMP is sufficient.
2. If the above is not applicable to the PI, they should confirm whether any investigators at their institution have received CMS approval of a DMP for a cloud-based storage environment. If possible, investigators should seek to re-use this environment for SEER-Medicare storage.
3. If neither option is available, investigators will need to complete a new DMP-SAQ and submit it to NCI for approval. The form can be found [here](https://healthcaredelivery.cancer.gov/seermedicare/obtain/SEER-Medicare_DMP_SAQ.docx).

Regardless of the storage method chosen, please provide the following directly after each bullet point:

* Physical address of the server where the data will be stored, and who will have access to it
* Describe how the data will be protected from unauthorized access (i.e. firewall)
* Provide information on the storage/protection of the physical media you receive containing the original files, including the address of the storage location and who will have access
* If an IT person is needed for uploading the data onto the servers, confirm that they have been added to the list of study personnel and have signed the custodian DUA
* Describe what you will do with the data at the end of the study (e.g., destruction of original media and data on server, notification of NCI)
* Provide assurance that no attempt will be made to identify patients, hospitals, or physicians
* Provide assurance that publications and presentations of the data will not allow identification of patients, hospitals, or physicians or publish cells where sizes are n<11
* Provide assurances that any publications or presentations of the data will not include any small geographic area maps (e.g., by zip code or census tract) without explicit approval of NCI and the involved registries

E. FUNDING SOURCE: If your organization is a consulting firm, contractor, biomedical or pharmaceutical company, or your project is funded by one of these entities, your application must include information about the funding source. This letter must come from a person in authority on company letterhead.

F. RESTRICTED DATA:

Selected data are not released without additional permission. If you are requesting access to any of the below mentioned data, you must include the justification in your description of the project and submit the completed request form for restricted variables. Please see <http://healthservices.cancer.gov/seermedicare/privacy/variables.html> for information.

1. Unencrypted census tract of the patient; zip code of the patient or provider; and unencrypted hospital provider numbers require special permission of the NCI and each SEER Registry. NCI will provide the requestor with contact information for each of the registries after approval at NCI; however, it is the responsibility of the requestor to obtain permission from each registry. A restricted variable request form is required.

**NOTE:** These unencrypted variables are not available from the Texas registry at this time.

1. Cases from the Alaska Native Tumor Registry have been linked to Medicare starting with the 2018 Linkage. These data are only available with special approval by the NCI and the Alaska Native Tumor Registry. NCI will provide the requestor with contact information for the Alaska Native Tumor Registry after approval at NCI. A restricted variable request form is required.
2. Applications requesting Oncotype Dx variables for breast cancer patients require special permission by the NCI. A restricted variable request form is required.

**NOTE:** All related applications and any manuscripts or reports that result from the analyses of such data will be shared with Genomic Health Inc (GHI), which is now part of Exact Sciences Corp., the company who developed the Oncotype Dx Assay. These documents will be shared with GHI for informational purposed only; all approval decisions will be handled by NCI.

1. Specialized SEER variables (e.g., head and neck HPV status, prostate watchful waiting) requires special approval by the NCI. Access to these variables requires is a separate approval process as outlined [here](https://seer.cancer.gov/data-software/specialized.html).
2. Applications requesting Area Deprivation Index (ADI) values require special permission by the NCI.

**NOTE:** All related applications and any manuscripts or reports that result including the ADI data will be shared with the ADI creators at the University of Wisconsin’s Neighborhood Atlas®.

1. Linkage to data about physicians from the American Medical Association (AMA) is feasible without release of the physician identifiers but requires special permission by the NCI and coordination with the AMA’s programming contractor, as described [here](https://healthcaredelivery.cancer.gov/seermedicare/privacy/variables.html).
2. Crosswalk between the Dartmouth Atlas on Health hospital referral regions (HRR) to encrypted zip codes requires special approval by the NCI.

**SECTION III. Data Files Requested:** Please list specific SEER-Medicare data files and years of data required. The project description **must** describe how each requested file will be used.

(Table on next page. Please complete the table with the years of the files being requested. Leave the ‘years requested’ cell blank for those files not requested.)

**These files will include:**

[ ]  Cancer cases [ ]  Non-cancer cases

| **Name of file** | **Years available** | **Years requested** |
| --- | --- | --- |
| **Cancer Data** |
| Cancer File | 1999-20191 |   |
| 5% Cancer File | 1999-20191 |   |
| **Medicare Enrollment** |
| Master Beneficiary Summary File (MBSF) Base (A/B/C/D)3 | 1999-2020 | 1999-20202 |
| Chronic Conditions Flags 27 conditions algorithm | 1999-2020 |   |
| Chronic Conditions Flags 30 conditions algorithm | 2017-2020 |  |
| Other Chronic or Potentially Disabling Conditions  | 2000-2020 |   |
| Plan Characteristics File  | 2007-2020 |   |
| **Medicare fee-for-service (FFS) Claims and Events Files** |
| MedPAR | 1999-2020 |   |
| Carrier Claims (NCH) | 1999-2020 |   |
| Outpatient | 1999-2020 |   |
| Home Health Agencies (HHA) | 1999-2020 |   |
| Hospice | 1999-2020 |   |
| Durable Medical Equipment (DME) | 1999-2020 |   |
| Part D Event (PDE)- with Drug Characteristics File appended | 2007-2020 |   |
| Formulary File | 2010-2020 |   |
| Prescriber Characteristics File and Bridge File | 2007-2020 |   |
| Pharmacy Characteristics File and Bridge File | 2007-2020 |   |
| Part D Medication Therapy Management File4 | 2013-2020 |   |
| **Medicare Advantage (MA) Encounter Data Files** |
| Inpatient | 2015-2019 |  |
| Skilled Nursing Facility | 2015-2019 |  |
| Carrier | 2015-2019 |  |
| Outpatient | 2015-2019 |  |
| Home Health Agencies (HHA) | 2015-2019 |  |
| Durable Medical Equipment (DME) | 2015-2019 |  |
| **Medicare Assessment Files5** |
| Minimum Data Set (MDS) | 1999-2020 |   |
| Outcome and Assessment Information Set (OASIS) | 1999-2020 |   |
| **Condensed Resources (CoRe) Files6**  |
|  CoRe Enrollment File | Dependent on Cancer File requested above and available Medicare data. Select the CoRe File(s) of interest.  | [ ]  |
|  CoRe Comorbidity- Prior to Cancer Diagnosis File | [ ]  |
|  CoRe Comorbidity- Post Cancer Diagnosis File | [ ]  |
|  CoRe Cancer Treatment- Systemic ABFFS File | [ ]  |
|  CoRe Cancer Treatment- Systemic Part D File | [ ]  |
|  CoRe Cancer Treatment- Radiation File | [ ]  |
| CoRe Cancer Treatment- Surgery File | [ ]  |
|  CoRe Cancer Treatment- Summary File | [ ]  |
| **Housing Assistance Data7** |
| Temporal Alignment File | 2006-2020 |  |
| Episode File | 2006-2020 |  |
| Transaction File | 2006-2020 |  |
| **Ancillary Files** |
| Medicare Data on Provider Practice and Specialty (MD-PPAS) | 2008-2020 |   |
| Hospital Characteristics File | 1996, 1998, 2000-2022 |   |
| Hospital Referral Regions (HHR) -zip code crosswalk8 | 1995-2019 |  |
| Geographic - zip code/census tract files (automatically provided) | 1999-2018 | 1999-2018 |

1Cancer cases from the expansion registries are diagnosed in 2000 or later.

2All years of MBSF enrollment information will automatically be provided for the requested cancer cases and, if applicable, for the non-cancer controls.

3Currently there is no 2019 data for non-cancer controls who were added to the Medicare 5% sample in 2019-2020; these data will be added as soon as possible.

4Currently there is no Part D MTM 2013-2019 data for cancer cases added in the LINK2022 (e.g., new cases diagnosed in 2018-2019); these data will be added as soon as possible.

5Persons added in LINK2022 (e.g., data made available in late 2022) will only have MDS and OASIS data from 2010+

6 The CoRe Files represent a summarization of the available Medicare enrollment and claims data among persons included in the specified Cancer File (diagnosis) years who met the CoRe cohort inclusion criteria: persons who are 66+ years old at malignant cancer diagnosis in 2000+ and were continuously enrolled in fee-for-service Parts A and B from 12 months prior to cancer diagnosis through at least one month post diagnosis and/or continuously enrolled in Part D from 4 months prior to cancer diagnosis through at least one month post diagnosis. Note: persons diagnosed at autopsy or death certificate were excluded, as are persons who died during the month of diagnosis.

7Limited to persons who were found in the SEER-Medicare data and also in the Department of Housing and Urban Development housing assistance data.

8Originally created in 2017 to include years 1995-2015. Updated in 2020 to include years 2016-2017 and in 2022 to include years 2018-2019.