

Investigator:	
Date	
Project title:	

**SEER-MEDICARE DATA USE AGREEMENT (DUA)
PRINCIPAL INVESTIGATOR**

Information pertaining to an individual's health status and medical treatment is sensitive. Therefore, specific laws, including the Privacy Act of 1974 and the Health Insurance Portability and Accountability Act of 1996, have been enacted to ensure the confidentiality of health information. In utilizing health data for research purposes, it is absolutely necessary to ensure, to the extent possible, that uses of such data will be limited to research. Uses for any other reason, particularly those resulting in personal disclosures, will be prosecuted to the full extent of the law. In addition, release of information about providers, i.e., the physicians and hospitals that provide care for cancer patients, may compromise the willingness of these providers to cooperate with the activities of the cancer registries. Therefore, considerations regarding the privacy of providers are also of great importance.

In order for the National Cancer Institute to provide the linked SEER-Surveillance, Epidemiology and End Results (SEER)-Medicare data to you, it is necessary that you agree to the following provisions:

1. You agree that the statements and methods made in your attached research proposal are complete and accurate.
2. You will not use the data for purposes other than described in your research proposal.
3. You will not permit others to use the data except for collaborators at your institution involved with the research as described in your proposal. Access to the SEER-Medicare data shall be limited to the minimum number of individuals necessary to achieve the purpose stated in your proposal. The specific location details of where the data will be stored must be provided in your proposal's data storage and management plan. If you plan to move the data to a new location at your institute you must contact NCI in writing prior to moving the data for instruction on how to handle the SEER-Medicare data.
4. You will establish and maintain the appropriate administrative, technical, and physical safeguards to protect the confidentiality of the data and to prevent unauthorized use or access to it, as described in your proposal. The safeguards shall provide a level and scope of security that is not less than the level and scope of security established by the Office of Management and Budget (OMB) in OMB Circular No. A-130, Appendix III—Security of Federal Automated Information Systems, which sets forth guidelines for security plans for automated information systems in Federal agencies.
5. You agree not to place the SEER-Medicare data on personal computers, portable devices and removable media without permission. Portable devices include any non-fixed equipment that contains an operating system which may be used to create, access or store SEER-Medicare data.

This includes but is not limited to laptops, personal digital assistants (PDAs), and smart phones. Removable media include, but are not limited to CDs, DVDs, MP3 players, removable memory, and USB drives (thumb drives). If approved, all data stored on any of these devices must be password protected AND encrypted. Approved encryption standards must be FIPS-140 compliant and include Advanced Encryption Algorithm (AES) that uses a 128, 192, or 256-bit key size. In the event that the data are lost or stolen, you agree to report the loss to the SEER-Medicare contact within 24-hours/first business day of discovering the loss. Cloud storage is only allowed after a Data Management Plan Self-Attestation Questionnaire (DMP-SAQ) has been approved.

6. You may use an institutionally provided VPN to link to a time-sharing system for data access. In this case, the remote PC may support the VPN, but the SEER-Medicare data must remain on the approved server. Additionally, access to the VPN shall be restricted to persons residing in the United States
7. You will store all media on which the SEER-Medicare data are delivered in a secure location, such as a locked file cabinet in a locked office, only accessible by you or appropriate designated staff.
8. You must maintain all datasets containing restricted variables physically separate from any other SEER-Medicare files. Separate access controls with strong user authentication (username/password, digital certifications, etc.) must be established to allow limited access to these files. You should be able to track all access to these files.
9. All SEER-Medicare data must reside at your institution under your purview. If you plan to leave this institution, you must contact NCI in writing prior to the transition for instructions on how to handle the SEER-Medicare data. You may not duplicate any SEER-Medicare files prior to the transition nor can you take SEER-Medicare data with you without written permission from NCI. All files under your purview must be destroyed prior to your departure or someone must agree to assume the responsibilities of the PI as described in this document.
10. You will not attempt to link nor permit others to link the SEER-Medicare data with individually identified records in another database without the written consent from NCI. The use of AI algorithms that combine multiple de-identified datasets to deduce an individual's identity is also prohibited.
11. No one having access to the data will attempt to learn the identity of any persons with cancer in these data and/or their physicians or treating hospitals. If you discover or are able to deduce the identity of a specific patient or provider (individual or institution), you agree that you will not attempt to contact these individuals or institutions.
12. I have read and understand the limitations of the SEER Radiation and chemotherapy data described on the [SEER website](#) and will include a description of relevant limitations in any analyses published using the SEER data. I acknowledge that NCI has substantive concerns about using the radiation and chemotherapy data to address certain research questions as described on the above website. I understand that any findings from such analyses may be inaccurate or misleading.
13. No findings or information derived from the SEER-Medicare data may be released if such findings contain any combination of data elements that might allow the deduction of a patient's or providers' (individual or institution) identity. Numbers less than 11 (eleven) must be suppressed.

Also, no use of percentages or other mathematical formulas may be used if they allow the derivation of patient, facility, or provider counts less than 11. Reporting of small area data (e.g., zip code, census tract, registry) derived from the SEER-Medicare files through maps, figures or tables is not permitted without prior approval from NCI. Mapping of data related to reflect incidence, treatment, or survival at the registry-specific level or at other small areas is not permitted without prior approval from NCI and the involved registries. Although it is permissible to report registry names with registry-specific cancer rates (e.g., incidence, complications, mortality), registry names must be anonymized when reporting the quality or completeness of registry-specific data (e.g., case or treatment ascertainment). You agree that NCI shall be the sole judge as to whether the anonymization sufficiently precludes one from identifying or deducing the identity of a specific patient, provider (individual or institution) or registry with a reasonable degree of certainty.

14. You agree to provide a copy of all manuscripts to NCI for review and comment prior to publication submission. You further agree not to submit such findings to any third party prior to completion of NCI review. NCI agrees to complete the manuscript review process within 4 weeks of receiving any manuscript. NCI's review of the manuscript is for the sole purpose of assuring that data confidentiality is maintained (e.g., individual patients and/or providers cannot be identified) and that the focus of the manuscript was outlined in the approved SEER-Medicare proposal. Revisions will be necessary, if NCI determines that the format in which data are presented may result in identification of individual patients and/or providers or if the scope of the manuscript is not consistent with the approved proposal.
15. If requesting Oncotype Dx data, you agree to allow NCI to share your application for SEER-Medicare data and any manuscripts or reports that result from the analyses of such data with Exact Sciences (formerly Genomic Health), the company that developed the Oncotype Dx Assay. These documents will be shared with GHI for informational purposes only; all approval processes will be handled by NCI
16. If receiving Area Deprivation Index (ADI) data, you agree to allow NCI to share information about your SEER-Medicare application and any resulting manuscripts with the creators of the ADI (Neighborhood Atlas, Center for Health Disparities Research, University of Wisconsin School of Medicine and Public Health). This data will be shared for informational purposes only; the Neighborhood Atlas monitors usage of the ADI data to justify continued financial support for these measures. You also agree to reference the Neighborhood Atlas citations, as listed in the SEER-Medicare Cancer File, in any resulting manuscript or report.
17. You agree that in the event NCI determines or has a reasonable belief that you have violated any terms of this agreement, NCI may request that you destroy the data and all derivative files and send a certificate/ notification of destruction to NCI. You understand that as a result of NCI's determination or reasonable belief that a violation of this agreement has taken place, NCI may refuse to release further SEER-Medicare data to you for a period of time to be determined by NCI.
18. All files received may be retained for a maximum of five years. At the completion of the project or five years from receipt all files including all back-up files and original media must be destroyed and notification of destruction must be sent to NCI. Investigators who need to retain files beyond that period must contact NCI.

On the following page, please indicate the SEER-Medicare files you will use and if the files will include:

Cancer cases

Non-cancer cases

Name of file	Years available	Years requested
Cancer Data		
Cancer File	1999-2021 ¹	
5% Cancer File	1999-2021 ¹	
Medicare Enrollment		
Master Beneficiary Summary File (MBSF) Base (A/B/C/D)	1999-2022	1999-2022 ²
Chronic Conditions Flags 27 conditions algorithm	1999-2021	
Chronic Conditions Flags 30 conditions algorithm	2017-2022	
Other Chronic or Potentially Disabling Conditions	2000-2022	
Plan Characteristics File	2007-2022	
Medicare fee-for-service (FFS) Claims and Events Files		
MedPAR	1999-2022	
Carrier Claims (NCH)	1999-2022	
Outpatient	1999-2022	
Home Health Agencies (HHA)	1999-2022	
Hospice	1999-2022	
Durable Medical Equipment (DME)	1999-2022	
Part D Event (PDE)- with Drug Characteristics File appended	2007-2022	
Formulary File	2010-2022	
Prescriber Characteristics File and Bridge File	2007-2022	
Pharmacy Characteristics File and Bridge File	2007-2022	
Part D Medication Therapy Management File	2013-2021	
Medicare Advantage (MA) Encounter Data Files		
Inpatient	2015-2021	
Skilled Nursing Facility	2015-2021	
Carrier	2015-2021	
Outpatient	2015-2021	
Home Health Agencies (HHA)	2015-2021	
Durable Medical Equipment (DME)	2015-2021	
Medicare Assessment Files ³		
Minimum Data Set (MDS)	1999-2022	
Outcome and Assessment Information Set (OASIS)	1999-2022	
Condensed Resources (CoRe) Files ⁴		FFS MA
CoRe Enrollment File- always with CoRe files, FFS/MA combined	Years vary by measure/data source.	
CoRe Comorbidity- Prior to Cancer Diagnosis File		
CoRe Comorbidity- Post Cancer Diagnosis File		
CoRe Cancer Treatment- Systemic ABFFS File / MA File	All available measures/file will be provided.	
CoRe Cancer Treatment- Systemic Part D File		
CoRe Cancer Treatment- Radiation File		
CoRe Cancer Treatment- Surgery File		
CoRe Cancer Treatment- Summary File		
Housing Assistance Data ⁵		
Temporal Alignment File	2006-2020	
Episode File	2006-2020	
Transaction File	2006-2020	
Ancillary Files		
Provider Characteristics		
Medicare Data on Provider Practice and Specialty (MD-PPAS)	2008-2022	
Hospital Characteristics File	1996, 1998, 2000-2022	

Name of file	Years available	Years requested
Hospital Referral Regions (HRR) -zip code crosswalk ⁶	1995-2019	
Area-level Files		Census Tract Zip Code
Traditional Census Files ⁷	1990, 2000	
New Files⁸		
Demographics & Employment Files	Years vary by measure/data source. All available options will be provided.	
Insurance & Health Status Files		
Housing & Mobility Files		

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.

³Persons added in LINK2022 and beyond will only have MDS and OASIS data from 2010+

⁴ 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.

⁵Limited to persons who were found in the SEER-Medicare data and in the Department of Housing and Urban Development housing assistance data.

⁶Originally created in 2017 to include year's 1995-2015. Updated in 2020 to include year's 2016-2017 and in 2022 to include year's 2018-2019.

⁷The Traditional Area Level Files (Census Tract & Zip Code) need to be specifically requested and approved. These files are free upon approval.

⁸First created in 2025, must select only one of the file types, to minimize small area re-identification; then select census tract and/or zip code.

Signature of Principal Investigator (In the case of students and fellows, the department chair or advisor from the student's academic institution must sign the data request)

Your signature indicates that you agree to comply with the above stated provisions. Deliberately making a false statement regarding any matter within the jurisdiction of any department or agency of the Federal Government violates 18 USC 1001 and is punishable by a fine up to \$10,000 or up to five years in prison.

Name – (printed or typed)
Institution/Organization
Street Address
City/State/ZIP code
Phone number – including Area Code
Email address
Signature
Date