がん登録推進法

国立がん研究センターがん対策情報センター
National Cancer Center
Center for Cancer Control and Information Services

- 第2章　全国がん登録
  第5節　情報の保護等　25-38条

全国がん登録等を取り扱うすべての者（職員等、研究者等の受領者）に対して

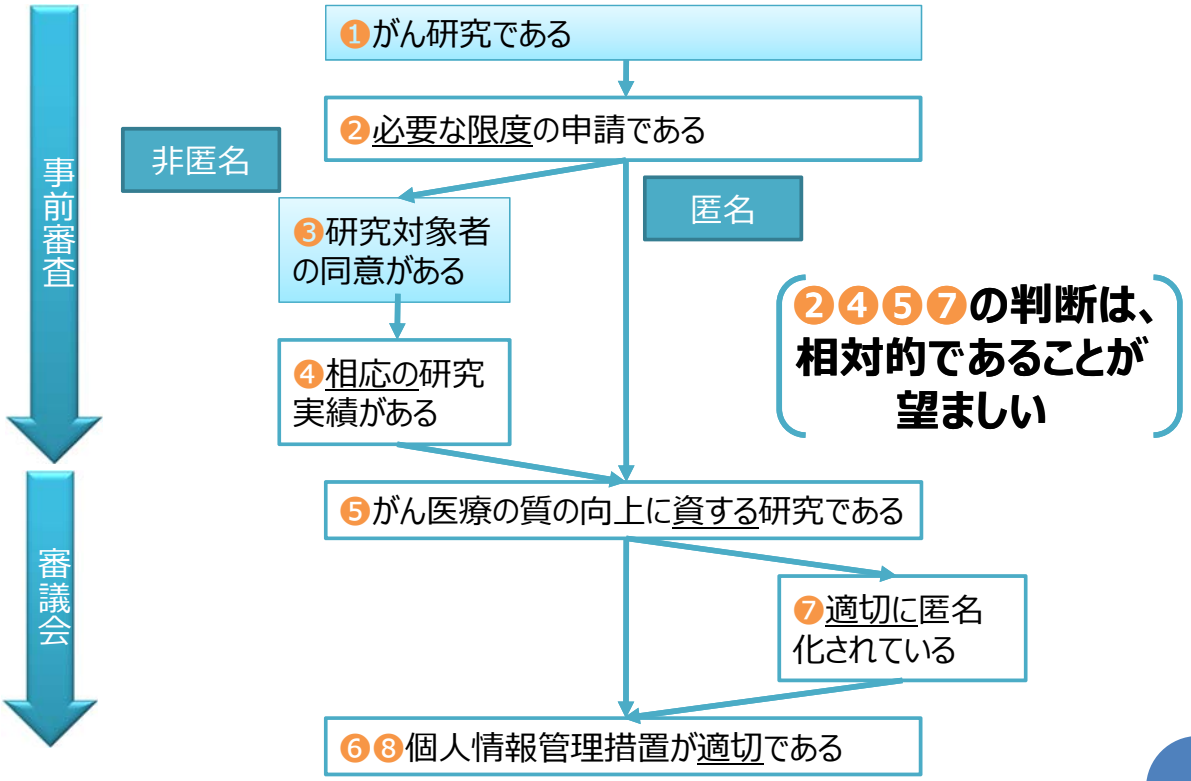
- 情報等の適切な管理
- 情報等の利用及び提供等の制限
- 情報等の保有等の制限（利用期間後の廃棄）
- 秘密保持義務、その他の義務

- 第2章　全国がん登録
  第3節　情報の利用及び提供　17-22条
  - 個人情報を適切に管理できる者又は個人を特定できない状態で提供

1

提供の前提条件と審議事項

国立がん研究センターがん対策情報センター
National Cancer Center
Center for Cancer Control and Information Services



2

提供の前提条件

非匿名化、匿名化情報共通

❶がんに係る調査研究であるか

❷申請内容は当該がんに係る調査研究に必要な限度を逸脱していないか

非匿名情報の提供

❸当該がんに罹患した者から、当該がんに係る調査研究のために当該全国がん登録情報が提供されることについて同意*を得ているか

❹申請者は、がんに係る調査研究であってがん医療の質の向上等に資するものの実績を相当程度有する者か

*同意：積極的同意（オプトイン同意）

3

審議会等

- 審議事項
  - 非匿名化情報、匿名化情報共通

    ❺当該がんに係る調査研究が、がん医療の質の向上等に資するものであるか
  - 非匿名化情報

    ❻秘密の漏えいの防止その他の当該がん全国がん登録情報の適切な管理のための必要な措置を講じているか
  - 匿名化情報

    ❼匿名化は適切か

    ❽当該匿名化が行われた情報について、その漏えい、減失及び毀損の防止その他の適切な管理のために必要な措置を講じているか

4

資料1

審議会等のあり方

「提供しない」ための審議会ではなく、提供する情報の限度、匿名化の適切さ、研究実績等の要素は、いかに『がん医療の質の向上に資する研究であるか』を基準に判断されることが望ましい

- 審議会の構成員
  - がん、がん医療等又はがんの予防に関する学識経験のある者及び個人情報の保護に関する学識経験のある者を含む
    - 全国がん登録情報の非匿名化情報の提供に関する審議会
      - 厚生科学審議会がん登録部会
    - 全国がん登録情報の匿名化及び匿名化情報の提供に関する審議会
      - 国がんに委任　法律第23条第1項、第2項の読替

【柴田メモ】平成29 年3 月　個人情報保護委員会
独立行政法人等の保有する個人情報の保護に関する法律についてのガイドライン
（独立行政法人等非識別加工情報編）
考慮の必要性について

5

法施行前に開始された研究への非匿名化情報の提供に係る経過措置

国立がん研究センターがん対策情報センター
National Cancer Center
Center for Cancer Control and Information Services

法施行後の必要条件
　当該がんに罹患した者から、当該がんに係る調査研究のために当該全国がん登録情報が提供されることについて**同意**を得ている

❶がん研究である

政令附則第2条第1項

❷H27.12.31以前に研究実施計画において対象者*の範囲が定められている

*対象者
・施行日前から対象とされている者
・施行日以後に、対象とされた者

政令附則第2条第2項

❸❷の対象者は、H28.1.1以後に「同意」を得ることが困難な事情のいずれかに当てはまるか

政令附則第2条第3項

❹研究対象者が5,000人以上

❹以下の事情**のため厚労大臣が同意取得困難と認定

**以下の事情
・施行日前から対象者とされている者と連絡をとることが困難
・改めて同意を取得することが研究の結果に影響を与える

政令附則第2条第4項申請

省令附則第2条

「同意」取得の代わりに、厚労大臣が指針で定める同意代替措置でよい

省令附則第2条

- 「事情があって改めて同意を取得することが困難である」ことについて、厚労大臣の認定を求める申請
  - 申請書の提出
    - 研究代表者の氏名、生年月日、住所
    - 研究の実施期間
    - 対象者の範囲及び数
    - 新たに同意取得することが困難である理由
  - 研究計画書の添付

告示　同意代替措置

調査研究を行う者が講ずる同意代替措置に関する指針

第一　調査研究を行う者が調査研究対象者に係る当該がんに係る情報を取得することについてのインフォームド・コンセントの取得等の実施

第二　調査研究を行う者が全国がん登録情報等の提供を受けることについての情報公開等の措置

一　適切な情報公開

二　調査研究対象者等が当該がんに係る調査研究のために全国がん登録情報等が提供されることについて拒否できる機会の保障

海外の動向

- EU: DIRECTIVE (EU) 2016/680 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA (2016年4月採択、2018年5月施行予定)
- US: Common ruleの改正（2016年3月）
  - 45 CFR 46 Code of Federal Regulations

    TITLE 45: PUBLIC WELFARE DEPARTMENT OF HEALTH AND HUMAN SERVICES

    PART 46: PROTECTION OF HUMAN SUBJECTS

US

FOR IMMEDIATE RELEASE
January 18, 2017
Contact: HHS Press Office
202-205-0143
ashmedia@hhs.gov

Final rule enhances protections for research participants, modernizes oversight system

Significant changes made in response to public comments

The U.S. Department of Health and Human Services and 15 other federal agencies today issued a final rule to update regulations that safeguard individuals who participate in research. Most provisions in the new rule will go into effect in 2018.

The new rule strengthens protections for people who volunteer to participate in research, while ensuring that the oversight system does not add inappropriate administrative burdens, particularly to low-risk research. It also allows more flexibility in keeping with today's dynamic research environment.

医療分野の研究開発に資するための匿名加工医療情報に関する法律案の概要（次世代医療基盤法案）

趣旨

特定の個人を識別できないように医療情報を匿名加工する事業者に対する規制を整備し、匿名加工された医療情報の安心・適正な利活用を通じて、健康・医療に関する先端的研究開発及び新産業創出を促進し、もって健康長寿社会の形成に資する。

概要

１．国の責務等

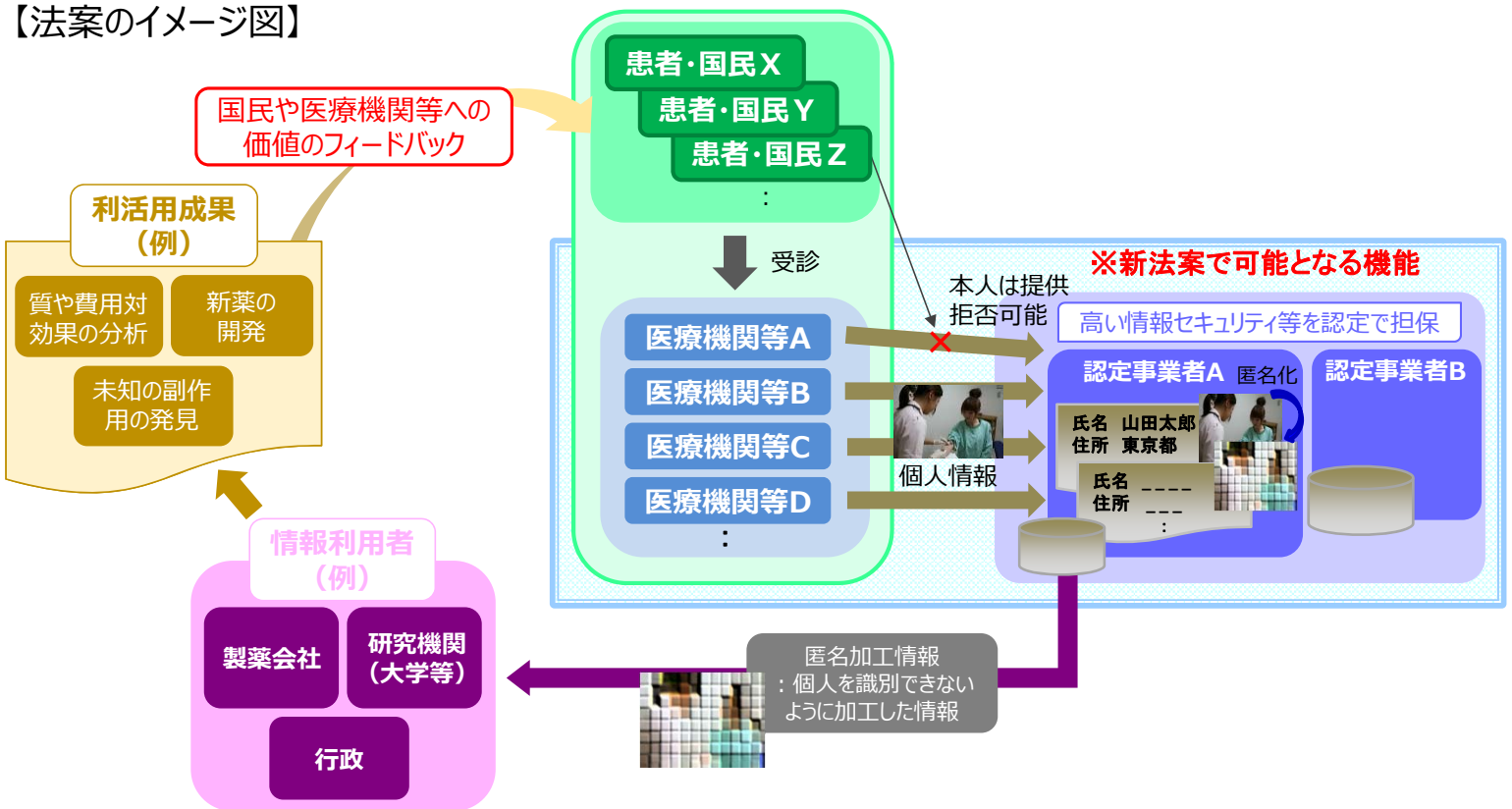
医療分野の研究開発に資するための匿名加工医療情報に関し、
（１）必要な施策を講ずる**国の責務**
（２）施策を総合的かつ一体的に推進するための**基本方針**　について定める。

２．認定匿名加工医療情報作成事業者（以下「認定事業者」という。）

（１）認定事業者の認定
高い情報セキュリティを確保し、十分な匿名加工技術を有するなどの**一定の基準**を満たし、医療情報等の管理や利活用のための匿名化を適正かつ確実に行うことができる者を**認定する仕組み**を設ける。

（２）医療情報等の取扱いに関する規制等
医療機関等は、あらかじめ本人に通知し、**本人が提供を拒否しない場合**、認定事業者に対し、**医療情報を提供できる**こととする。
（医療機関等から認定事業者への医療情報の提供は任意）

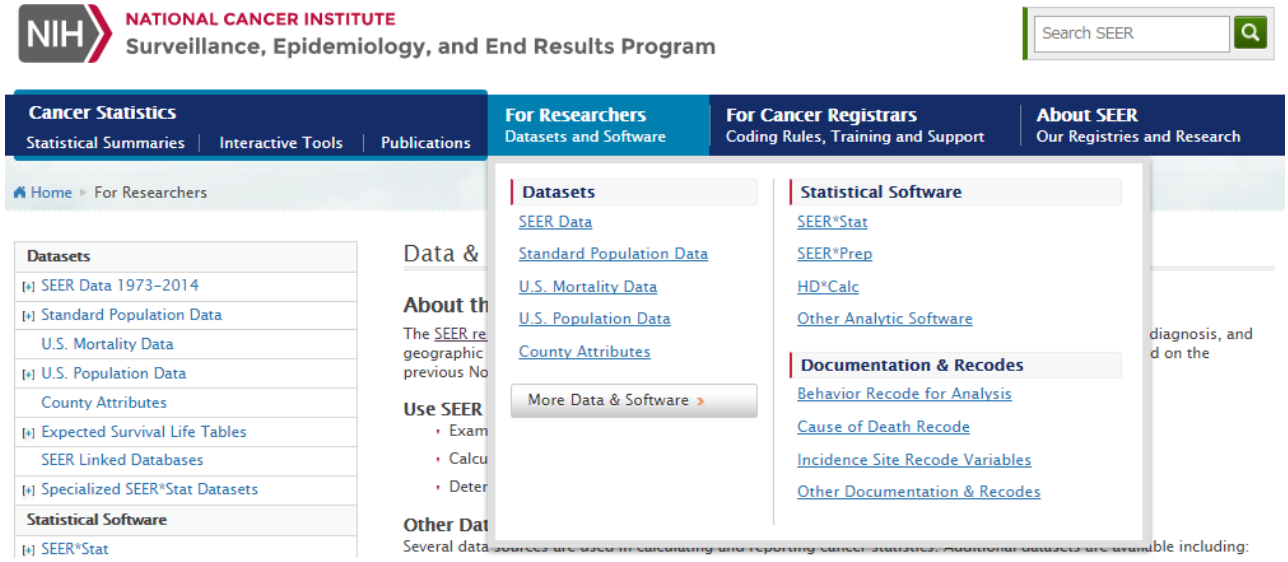
【法案のイメージ図】



SEER site：Accessing the 1973-2014 SEER Data の説明

国立がん研究センターがん対策情報センター
がん統計・総合解析研究部　堀 芽久美

SEER ホームページにおいて、SEER Data の利用申請、インターネットを経由した SEER Data へのアクセス、SEER Data 集計ができるソフトウェア SEER*Stat のダウンロードが可能



A) SEER Data 1973-2014 へのアクセス

a) 取得できるデータ

- ファイル形式：SEER*Stat 用のバイナリ形式ファイル、ASCII テキスト形式
- アクセス方法：①SEER*Stat を利用したインターネット経由（SEER*Stat's client-server mode）でアクセス、②圧縮ファイルをダウンロード、③DVD の郵送
- 含まれるデータセット、アクセス方法別：

データセット	アクセス方法		
	①	②	③
SEER Research Data, 1973-2014 (9, 13, and 18 registries databases)	○	○	○
County Attributes data	○	○	○
US Mortality	○		
Previous submissions of SEER data	○		
Incidence with SEER Delay Factors	○		
US Populations	○		
Specialized SEER*Stat Datasets	○		
Multiple primaries-standardized mortality ratios (MP-SMRs)	○		
Automatic SEER*Stat software updates	○		
ASCII text version of the SEER 1973-2014 data		○	

b) データ取得までの手順

・ オンライン利用申請（申請書提出→ID 取得→誓約書提出→承認・パスワード通知）

» 申請内容：アクセス方法（インターネット経由、DVD 郵送)、当該データに対する過去の利用申請および誓約書提出の有無（図 1)、申請者の情報（図 2）

・ 申請書の提出後、個別の誓約書フォーマットの URL がメールで知らされるため、署名し、pdf を添付して指定のアドレスに返信（図 3）

» 誓約書の内容

1) 研究目的以外で利用しない
2) 個人が特定されるデータを公表しない、少数例の集計値の公表を避ける
3) 他のデータベースの個人単位のレコードとのリンケージをしない
4) 個人を特定しようとしない
5) 個人が特定された場合には、その情報を使用しない、SEER にインシデント報告、特定された個人を知らせない
6) SEER の承認を得た人以外にデータを利用させない
7) 適切なデータ保管措置を行う
8) SEER から提供されるソフトウェアの適切な使用
9) 結果公表の際の適切な引用

・ 承認されると SEER*Stat's client-server へのログインパスワードがメール送信される

B) SEER*Stat を利用したインターネット経由での利用

a) 手順（Session → Database → (Statistic), Selection, Table, Output）

・ Session：Frequency, Rate, Survival, Prevalence, MP-SIR, Life-Table, Case Listing から選択

・ Database：Database の選択

データベース名
Incidence- SEER 18 Regs Research Data + …, Nov 2016 Sub (1973-2014 varying)
Incidence- SEER 18 Regs Research Data + …, Nov 2015 Sub (1973-2013 varying)
・・・
Incidence- SEER 9 Regs Research Data, Nov 2003 Sub (1973-2011)
・・・
Mortality- All COD, Aggregated With State, Total U.S. (1969-2014) …
・・・
Populations- Total U.S. (1969-2015) …
・・・
County Attributes- Total U.S., 1969-2015 Counties.
・・・

・ Statistic：集計方法の設定

・ Selections：集計対象の指定

・ Table：集計・抽出する変数の選択

b) 出力例

・図 4　卵巣がんの組織/性状別、年齢階級別罹患数

Session	: Frequency Session
Database	: Incidence- SEER 18 Regs Research Data + …, Nov 2016 Sub (1973-2014 varying)
Statistic	: Statistic=Frequencies, Percentage=None
Selection	: Year of diagnosis = ”2014”
Table	: Row= Age, Colum= ICD-O-3 Hist/behave

・図 5　卵巣がん症例リスト

Session	: Case Listing Session
Database	: Incidence- SEER 18 Regs Research Data + …, Nov 2016 Sub (1973-2014 varying)
Selection	: Year of diagnosis = ”2014”
Table	: Column=Age, Race, ICD-O-3 Hist/behave

c) データセットに含まれる項目（別添 1）

C) SEER Linked Databases へのアクセス

a) SEER-Medicare Linked Database

- SEER Data に含まれる人と、その Medicare 請求を NCI と Centers for Medicare and Medicaid Services（CMS）がリンケージしたデータ
- 個人特定される項目は削除・暗号化　* 非暗号化項目は必要な場合に申請可
- 保険受給者データからランダムに 5%抽出した非がん患者を含む

b) データ取得までの手順

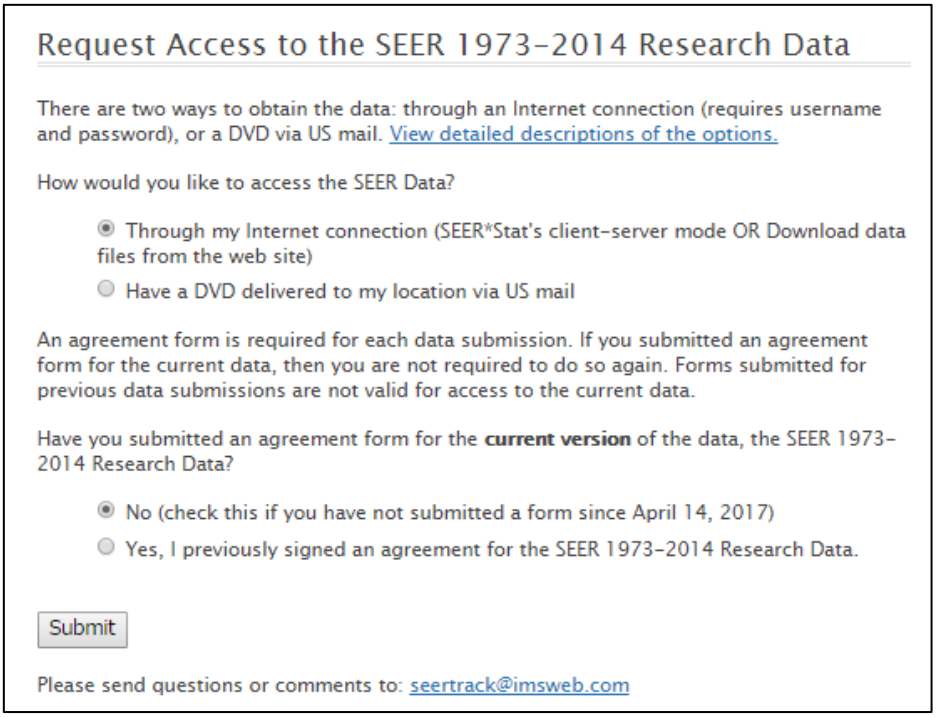
- 以下の書類を添付してメール送付（NCI の SEER-Medicare 窓口へ）（別添 2）
  - » 申請書の提出（研究計画の概要）
  - » 誓約書
  - » 所属機関の倫理審査委員会の承認書
  - » （必要な場合のみ）特定の項目（非暗号化の患者や医療機関・保険者の郵便番号、患者の人口調査標準地域、医療機関コード）の利用申請書　* SEER の各登録からの承認が必要
- 承認審査は 4～6 週間、有料（図 6）
- NCI と SEER の代表者が承認審査を行う

（SEER 参加登録による調査・研究の場合）

- » SEER の研究責任者からのレター　*SEER 登録室主体の研究であること、その研究の協同研究者であることを示す内容
- » 誓約書
- » 所属機関の倫理審査委員会の承認書

- 年 4 件まで無料

図 1



Request Access to the SEER 1973–2014 Research Data

There are two ways to obtain the data: through an Internet connection (requires username and password), or a DVD via US mail. View detailed descriptions of the options.

How would you like to access the SEER Data?

- (•) Through my Internet connection (SEER*Stat's client–server mode OR Download data files from the web site)
- ( ) Have a DVD delivered to my location via US mail

An agreement form is required for each data submission. If you submitted an agreement form for the current data, then you are not required to do so again. Forms submitted for previous data submissions are not valid for access to the current data.

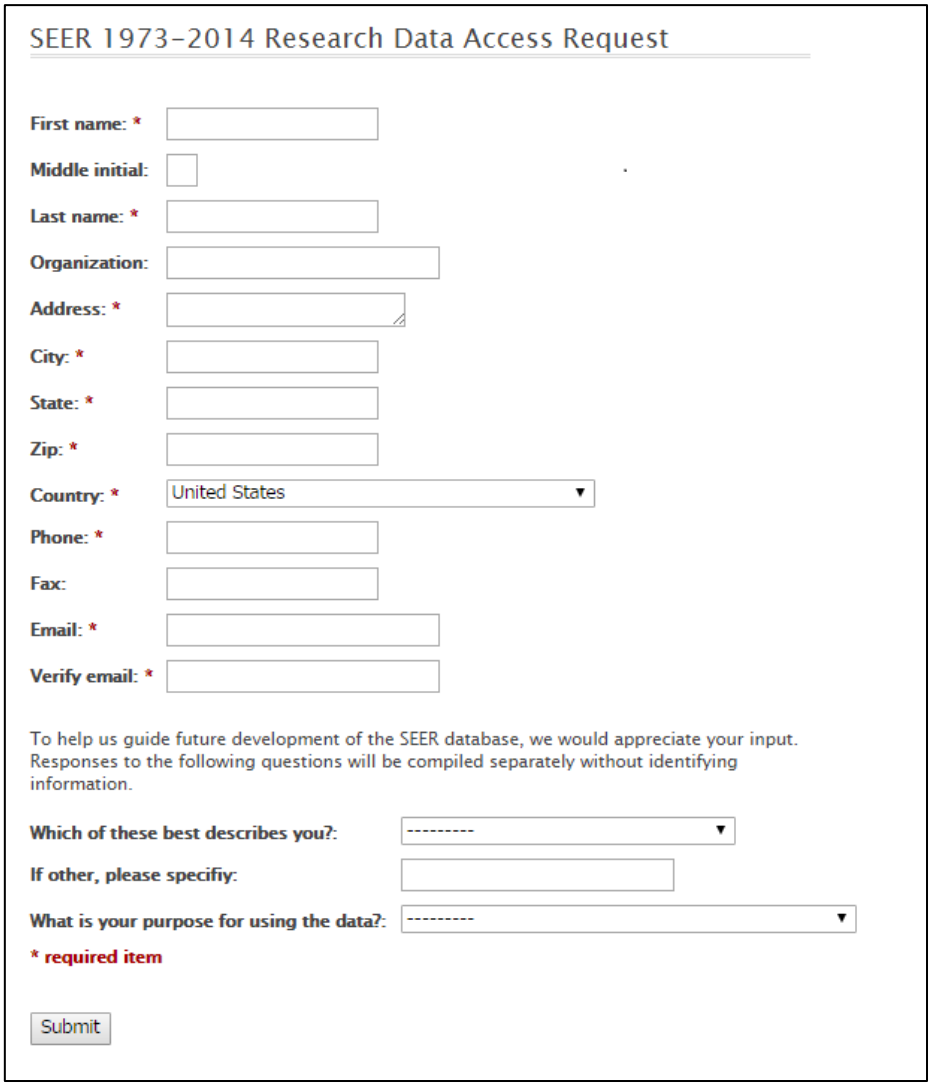
Have you submitted an agreement form for the **current version** of the data, the SEER 1973–2014 Research Data?

- (•) No (check this if you have not submitted a form since April 14, 2017)
- ( ) Yes, I previously signed an agreement for the SEER 1973–2014 Research Data.

Submit

Please send questions or comments to: seertrack@imsweb.com

図 2



SEER 1973–2014 Research Data Access Request

First name: *
Middle initial:
Last name: *
Organization:
Address: *
City: *
State: *
Zip: *
Country: * United States
Phone: *
Fax:
Email: *
Verify email: *

To help us guide future development of the SEER database, we would appreciate your input. Responses to the following questions will be compiled separately without identifying information.

Which of these best describes you?: ---------
If other, please specifiy:
What is your purpose for using the data?: ---------

*** required item**

Submit

図 3

Last Name: Hori
SEER ID: 11505-Nov2016
Request Type: Internet Access

SURVEILLANCE, EPIDEMIOLOGY, AND END RESULTS PROGRAM
Data-Use Agreement for the SEER 1973-2014 Research Data File

It is of utmost importance to protect the identities of cancer patients. Every effort has been made to exclude identifying information on individual patients from the computer files. Certain demographic information - such as sex, race, etc. - has been included for research purposes. All research results must be presented or published in a manner that ensures that no individual can be identified. In addition, there must be no attempt either to identify individuals from any computer file or to link with a computer file containing patient identifiers.

In order for the Surveillance, Epidemiology, and End Results Program to provide access to its Research Data File to you, it is necessary that you agree to the following provisions.

1. I will not use - or permit others to use - the data in any way other than for statistical reporting and analysis for research purposes. I must notify the SEER Program if I discover that there has been any other use of the data.
2. I will not present or publish data in which an individual patient can be identified. I will not publish any information on an individual patient, including any information generated on an individual case by the case listing session of SEER*Stat. In addition, I will avoid publication of statistics for very small groups.
3. I will not attempt either to link - or permit others to link - the data with individual level records in another database.
4. I will not attempt to learn the identity of any patient whose cancer data is contained in the supplied file(s).
5. If I inadvertently discover the identity of any patient, then (a) I will make no use of this knowledge, (b) I will notify the SEER Program of the incident, and (c) I will inform no one else of the discovered identity.
6. I will not either release - or permit others to release - the data - in full or in part - to any person except with the written approval of the SEER Program. In particular, all members of a research team who have access to the data must sign this data-use agreement.
7. I will use appropriate safeguards to prevent use or disclosure of the information other than as provided for by this data-use agreement. If accessing the data from a centralized location on a time sharing computer system or LAN with SEER*Stat or another statistical package, I will not share my logon name or password with any other individuals. I will also not allow any other individuals to use my computer account after I have logged on with my logon name and password.
8. For all software provided by the SEER Program, I will not copy it, distribute it, reverse engineer it, profit from its sale or use, or incorporate it in any other software system.
9. I will cite the source of information in all publications. The appropriate citation is associated with the data file used. (Please see either Suggested Citations on the SEER*Stat Help menu or the Readme.txt associated with the ASCII text version of the SEER data.)

My signature indicates that I agree to comply with the above stated provisions.

_____
Signature

_____
Date

Please print, sign, and date the agreement. Send the form to The SEER Program:

- **By fax to 301-680-9571**
- **Or, e-mail a scanned form to seerfax@imsweb.com**

Last Name: Hori | SEER ID: 11505-Nov2016 | Request Type: Internet Access

図 4

SEER*Stat 8.3.4 - [Frequency Session-3 Matrix-2]

File Edit Matrix Window Profile Help

	00 years	01-04 years	05-09 years	10-14 years	15-19 years	20-24 years	25-29
8000/0: Neoplasm, benign	0	0	0	0	0	0	
8000/1: Neoplasm, uncertain whether benign or malignant	0	0	0	0	0	0	
8000/2: Neoplasm, in situ	0	0	0	0	0	0	
8000/3: Neoplasm, malignant	1	2	2	3	4	21	
8001/0: Tumor cells, benign	0	0	0	0	0	0	
8001/1: Tumor cells, uncertain whether benign or malignant	0	0	0	0	0	0	
8001/3: Tumor cells, malignant	0	0	1	0	0	0	
8002/3: Malignant tumor, small cell type	0	0	0	0	0	0	
8003/3: Malignant tumor, giant cell type	0	0	0	0	0	0	
8004/3: Malignant tumor, spindle cell type	0	0	0	2	0	0	
8005/3: Malignant tumor, clear cell type	0	0	0	1	0	0	
8010/0: Epithelial tumor, benign	0	0	0	0	0	0	
8010/2: Carcinoma in situ, NOS	0	0	0	0	0	0	
8010/3: Carcinoma, NOS	0	0	0	1	1	12	
8011/2: Epithelioma, in situ	0	0	0	0	0	0	
8011/3: Epithelioma, malignant	0	0	0	0	0	0	
8012/2: Large cell carcinoma in situ	0	0	0	0	0	0	
8012/3: Large cell carcinoma, NOS	0	0	0	0	0	0	
8013/3: Large cell neuroendocrine carcinoma	0	0	0	0	0	0	

図 5

Case Listing Session-1 Matrix-1

	Age recode with <1 year olds	Race recode (White, Black, Other)	ICD-O-3 Hist/behav
1	85+ years	White	9732/3: Plasma cell myeloma
2	55-59 years	White	8500/3: Infiltrating duct carcinoma, NOS
3	65-69 years	Other (American Indian/AK Native, Asian/Pacific Islander)	8140/3: Adenocarcinoma, NOS
4	65-69 years	Black	8041/3: Small cell carcinoma, NOS
5	55-59 years	White	8144/3: Adenocarcinoma, intestinal type
6	25-29 years	White	9401/3: Astrocytoma, anaplastic
7	85+ years	White	8140/3: Adenocarcinoma, NOS
8	45-49 years	White	8050/3: Papillary carcinoma, NOS
9	85+ years	White	8410/3: Sebaceous adenocarcinoma
10	55-59 years	White	9823/3: Chronic lymphocytic leukemia/small lymphocytic lymphoma
11	60-64 years	Black	8041/3: Small cell carcinoma, NOS
12	85+ years	Other (American Indian/AK Native, Asian/Pacific Islander)	8140/3: Adenocarcinoma, NOS
13	75-79 years	White	8720/3: Malignant melanoma, NOS
14	50-54 years	White	8310/3: Clear cell adenocarcinoma, NOS
15	85+ years	White	8500/3: Infiltrating duct carcinoma, NOS
16	85+ years	White	8070/3: Squamous cell carcinoma, NOS
17	85+ years	White	8720/3: Malignant melanoma, NOS
18	65-69 years	Black	8560/3: Adenosquamous carcinoma
19	65-69 years	White	8041/3: Small cell carcinoma, NOS
20	85+ years	Black	8140/3: Adenocarcinoma, NOS
21	45-49 years	White	8500/3: Infiltrating duct carcinoma, NOS

図 6

SEER-Medicare: Cost of Acquiring Data

The SEER-Medicare data files are created by Information Management Services, Inc. (IMS), the information technology contractor for NCI. Outside researchers are asked to reimburse IMS for the cost of producing the data, as listed below.

The quoted costs are for files containing cancer cases that are newly diagnosed in each year. Other analyses, such as identifying all persons who died of a selected cancer in a year, may require special programming at additional cost. Please note that IMS cannot undertake extensive special programming. Additional cleaning and data modification are the responsibility of the investigator. IMS will begin processing a data request upon receipt of payment. Completed datasets will be shipped by UPS ground.

All charges will be periodically reviewed to determine whether they are fairly recovering IMS's costs and may be changed without notice.

	Cost of Cancer Data File (including Data Production & DVD Media Charges)	Cost of Non-cancer Data File (including Data Production & DVD Media Charges)
Entitlement Data		
Cancer Cases: Patient Entitlement and Diagnosis Summary File (PEDSF)[1]	$110	N/A
Non-cancer Cases: Summarized Denominator File (SUMDENOM)[2]	N/A	$60
Claims Files		
Medicare Provider Analysis and Review (MEDPAR)	$110	$60
Carrier Claims (NCH)	$85/year	$210
Outpatient (OUTPAT)	$60/year	$210
Home Health Agencies (HHA)	$260	$110
Hospice	$260	$110
Durable Medical Equipment (DME)	$260	$110
Part D Event (PDE)	$60/year	$210
Chronic Conditions Flags	$110	$60

1: The PEDSF file must be ordered.
2: The SUMDENOM file must be ordered when ordering any non-cancer claim files.

These estimates are for creating data files for one cancer site, such as lung. Use the charges in the above table to estimate the total cost of your request. The estimated costs for the data files can also be calculated on this Web site.

SEER RESEARCH DATA RECORD DESCRIPTION

CASES DIAGNOSED IN 1973-2014*

Submission: *November 2016*

Follow-up Cutoff Date: *December 31, 2014*

Documentation Version: *April 2017*

Diagnosis Years: *1973-2014*

* *This documentation describes the data files in the incidence/yr1973_2014.seer9, yr1992_2014.sj_la_rg_ak, yr2000_2014.ca_ky_lo_nj_ga , and yr2005.lo_2nd_half directories. Refer to individual variable definitions to determine the differences between the directory files.*

Cervix in situ cases after 1995 are not included.

COMPUTER RECORD FORMAT

NAACCR Name	NAACCR Item #	SAS Variable Name	Applicable Years	Position	Length
Patient ID number	20	PUBCSNUM		1-8	8
Registry ID	40	REG		9-18	10
Marital Status at DX	150	MAR_STAT		19	1
Race/Ethnicity	160	RACE1V		20-21	2
NHIA Derived Hispanic Origin	191	NHIADE		23	1
Sex	220	SEX		24	1
Age at diagnosis	230	AGE_DX		25-27	3
Year of Birth	240	YR_BRTH		28-31	4
Sequence Number—Central	380	SEQ_NUM		35-36	2
Month of diagnosis	390	MDXRECMP		37-38	2
Year of diagnosis	390	YEAR_DX		39-42	4
Primary Site	400	PRIMSITE		43-46	4
Laterality	410	LATERAL		47	1
Histology (92-00) ICD-O-2	420	HISTO2V		48-51	4
Behavior (92-00) ICD-O-2	430	BEHO2V		52	1
Histologic Type ICD-O-3	522	HISTO3V		53-56	4
Behavior Code ICD-O-3	523	BEHO3V		57	1
Grade	440	GRADE		58	1
Diagnostic Confirmation	490	DX_CONF		59	1
Type of Reporting Source	500	REPT_SRC		60	1
EOD—Tumor Size	780	EOD10_SZ	1988-2003	61-63	3
EOD—Extension	790	EOD10_EX	1988-2003	64-65	2
EOD—Extension Prost Path	800	EOD10_PE	1985-2003	66-67	2
EOD—Lymph Node Involv	810	EOD10_ND	1988-2003	68	1
Regional Nodes Positive	820	EOD10_PN	1988+	69-70	2
Regional Nodes Examined	830	EOD10_NE	1988+	71-72	2
EOD—Old 13 Digit	840	EOD13	1973-1982	73-85	13
EOD—Old 2 Digit	850	EOD2	1973-1982	86-87	2
EOD—Old 4 Digit	860	EOD4	1983-1987	88-91	4
Coding System for EOD	870	EOD_CODE	1973-2003	92	1
Tumor Marker 1	1150	TUMOR_1V	1990-2003	93	1
Tumor Marker 2	1160	TUMOR_2V	1990-2003	94	1
Tumor Marker 3	1170	TUMOR_3V	1998-2003	95	1
CS Tumor Size	2800	CSTUMSIZ	2004+	96-98	3
CS Extension	2810	CSEXTEN	2004+	99-101	3
CS Lymph Nodes	2830	CSLYMPHN	2004+	102-104	3
CS Mets at Dx	2850	CSMETSDX	2004+	105-106	2
CS Site-Specific Factor 1	2880	CS1SITE	2004+*	107-109	3
CS Site-Specific Factor 2	2890	CS2SITE	2004+*	110-112	3
CS Site-Specific Factor 3	2900	CS3SITE	2004+*	113-115	3
CS Site-Specific Factor 4	2910	CS4SITE	2004+*	116-118	3
CS Site-Specific Factor 5	2920	CS5SITE	2004+*	119-121	3

NAACCR Name	NAACCR Item #	SAS Variable Name	Applicable Years	Position	Length
CS Site-Specific Factor 6	2930	CS6SITE	2004+*	122-124	3
CS Site-Specific Factor 25	2879	CS25SITE	2004+*	125-127	3
Derived AJCC T	2940	DAJCCT	2004+	128-129	2
Derived AJCC N	2960	DAJCCN	2004+	130-131	2
Derived AJCC M	2980	DAJCCM	2004+	132-133	2
Derived AJCC Stage Group	3000	DAJCCSTG	2004+	134-135	2
Derived SS1977	3010	DSS1977S	2004+	136	1
Derived SS2000	3020	DSS2000S	2004+	137	1
Derived AJCC—Flag	3030	DAJCCFL	2004+	138	1
CS Version Input Original	2935	CSVFIRST	2004+	141-146	6
CS Version Derived	2936	CSVLATES	2004+	147-152	6
CS Version Input Current	2937	CSVCURRENT	2004+	153-158	6
RX Summ—Surg Prim Site	1290	SURGPRIF	1998+	159-160	2
RX Summ—Scope Reg LN Sur	1292	SURGSCOF	2003+	161	1
RX Summ—Surg Oth Reg/Dis	1294	SURGSITF	2003+	162	1
RX Summ—Reg LN Examined	1296	NUMNODES	1998-2002	163-164	2
Reason for no surgery	1340	NO_SURG		166	1
RX Summ—Surgery Type	1640	SS_SURG	1973-1997	170-171	2
RX Summ—Scope Reg 98-02	1647	SURGSCOP	1998-2002	174	1
RX Summ—Surg Oth 98-02	1648	SURGSITE	1998-2002	175	1
SEER Record Number	2190	REC_NO		176-177	2
SEER Type of Follow-up	2180	TYPE_FU		191	1
Age Recode <1 Year olds	N/A	AGE_1REC		192-193	2
Site Recode ICD-O-3/WHO 2008	N/A	SITERWHO		199-203	5
Recode ICD-O-2 to 9	N/A	ICDOTO9V		204-207	4
Recode ICD-O-2 to 10	N/A	ICDOT10V		208-211	4
ICCC site recode ICD-O-3/WHO 2008	N/A	ICCC3WHO		218-220	3
ICCC site rec extended ICD-O-3/WHO 2008	N/A	ICCC3XWHO		221-223	3
Behavior Recode for Analysis	N/A	BEHTREND		224	1
Histology Recode—Broad Groupings	N/A	HISTREC		226-227	2
Histology Recode—Brain Groupings	N/A	HISTRECB		228-229	2
CS Schema v0204+	N/A	cs0204schema		230-232	3
Race recode (White, Black, Other)	N/A	RAC_RECA		233	1
Race recode (W, B, AI, API)	N/A	RAC_RECY		234	1
Origin recode NHIA (Hispanic, Non-Hisp)	N/A	ORIGRECB		235	1
SEER historic stage A	N/A	HST_STGA		236	1
AJCC stage 3rd edition (1988-2003)	N/A	AJCC_STG		237-238	2
SEER modified AJCC Stage 3rd ed (1988-2003)	N/A	AJ_3SEER		239-240	2
SEER Summary Stage 1977 (1995-2000)	N/A	SSS77VZ	1995-2000	241	1
SEER Summary Stage 2000 (2001-2003)	N/A	SSSM2KPZ	2001-2003	242	1

NAACCR Name	NAACCR Item #	SAS Variable Name	Applicable Years	Position	Length
First malignant primary indicator	N/A	FIRSTPRM		245	1
State-county recode	N/A	ST_CNTY		246-250	5
Cause of Death to SEER site recode	N/A	CODPUB		255-259	5
COD to site rec KM	N/A	CODPUBKM		260-264	5
Vital Status recode	N/A	STAT_REC		265	1
IHS Link	192	IHSLINK		266	1
Summary stage 2000 (1998+)	N/A	SUMM2K	1998+	267	1
AYA site recode/WHO 2008	N/A	AYASITERWHO		268-269	2
Lymphoma subtype recode/WHO 2008	N/A	LYMSUBRWHO		270-271	2
SEER Cause-Specific Death Classification	N/A	VSRTSADX		272	1
SEER Other Cause of Death Classification	N/A	ODTHCLASS		273	1
CS Tumor Size/Ext Eval	2820	CSTSEVAL	2004+	274	1
CS Lymph Nodes Eval	2840	CSRGEVAL	2004+	275	1
CS Mets Eval	2860	CSMTEVAL	2004+	276	1
Primary by international rules	N/A	intprim		277	1
ER Status Recode Breast Cancer (1990+)	N/A	erstatus	1990+	278	1
PR Status Recode Breast Cancer (1990+)	N/A	prstatus	1990+	279	1
CS Schema -AJCC 6th ed (previously called v1)	N/A	csschema		280-281	2
CS Site-Specific Factor 8	2862	CS8SITE	2004+*	282-284	3
CS Site-Specific Factor 10	2864	CS10SITE	2004+*	285-287	3
CS Site-Specific Factor 11	2865	CS11SITE	2004+*	288-290	3
CS Site-Specific Factor 13	2867	CS13SITE	2004+*	291-293	3
CS Site-Specific Factor 15	2869	CS15SITE	2004+*	294-296	3
CS Site-Specific Factor 16	2870	CS16SITE	2004+*	297-299	3
Lymph vascular invasion	1182	VASINV	2004+*	300	1
Survival months	N/A	srv_time_mon		301-304	4
Survival months flag	N/A	srv_time_mon_flag		305	1
Insurance recode (2007+)	N/A	INSREC_PUB	2007+	311	1
Derived AJCC-7 T	3400	DAJCC7T	2010+	312-314	3
Derived AJCC-7 N	3410	DAJCC7N	2010+	315-317	3
Derived AJCC-7 M	3420	DAJCC7M	2010+	318-320	3
Derived AJCC-7 Stage Grp	3430	DAJCC7STG	2010+	321-323	3
Breast Adjusted AJCC 6th T (1988+)	N/A	ADJTM_6VALUE	1988+	324-325	2
Breast Adjusted AJCC 6th N (1988+)	N/A	ADJNM_6VALUE	1988+	326-327	2
Breast Adjusted AJCC 6th M (1988+)	N/A	ADJM_6VALUE	1988+	328-329	2
Breast Adjusted AJCC 6th Stage (1988+)	N/A	ADJAJCCSTG	1988+	330-331	2
CS Site-Specific Factor 7	2861	CS7SITE	2004+*	332-334	3
CS Site-Specific Factor 9	2863	CS9SITE	2004+*	335-337	3
CS Site-Specific Factor 12	2866	CS12SITE	2004+*	338-340	3
Derived HER2 Recode (2010+)	N/A	her2	2010+	341	1

NAACCR Name	NAACCR Item #	SAS Variable Name	Applicable Years	Position	Length
Breast Subtype (2010+)	N/A	brst_sub	2010+	342	1
Lymphomas: Ann Arbor Staging (1983+)	N/A	ANNARBOR	1983+	348	1
CS Mets at Dx-Bone	2851	CSMETSDXB_PUB	2010+	349	1
CS Mets at Dx-Brain	2852	CSMETSDXBR_PUB	2010+	350	1
CS Mets at Dx-Liver	2853	CSMETSDXLIV_PUB	2010+	351	1
CS Mets at Dx-Lung	2854	CSMETSDXLUNG_PUB	2010+	352	1
T value - based on AJCC 3rd (1988-2003)	N/A	T_VALUE	1988-2003	353	2
N value - based on AJCC 3rd (1988-2003)	N/A	N_VALUE	1988-2003	355	2
M value - based on AJCC 3rd (1988-2003)	N/A	M_VALUE	1988-2003	357	2
Total Number of In Situ/malignant Tumors for Patient	N/A	MALIGCOUNT		359-360	2
Total Number of Benign/Borderline Tumors for Patient	N/A	BENBORDCOUNT		361-362	2

*2004+ varying by schema

Dear Investigator

Thank you for your interest in the SEER-Medicare data. Please use this application form to request SEER-Medicare files from the SEER-Medicare 2016 linkage. In order to facilitate the review process, you must complete and provide all items ***on this form***. Do not say "see attached". Be sure to review and include all required elements as listed in the application checklist. Incomplete applications will be delayed. You must submit this completed form ***electronically*** to the SEER-Medicare contact along a completed and signed Data Use Agreement (DUA), documentation of IRB approval and, if necessary the request for restricted variable form. Instructions for submitting a request can be found at http://healthservices.cancer.gov/seermedicare/obtain/requests.html .

Thank you

NCI SEER-Medicare Staff

Questions should be sent to the SEER Medicare contact:

Elaine Yanisko
yaniskoe@imsweb.com

Application Checklist (for your use only)

To be sent by email attachment to the SEER-Medicare contact:

☐ Application:

☐ Your description of the project **must** include:
- statement of main hypothesis / specific research question
- description of study subjects and cancer sites/phases to be included in the analysis
- brief explanation of how key components of the study will be obtained/identified within the PEDSF and/or claims data– specifically:
  - cohort selection criteria
  - covariates
  - outcomes
- a list of requested files and how each will be used, (for example: MEDPAR will be used …. NCH will be used….).
- how the 5% population (non-cancer and/or other cancer) will be used, if requested
- description of the personnel involved
- timeline for completion
- references can be included, if relevant.

☐ You **must** include an explanation of data storage and protection. Please **be specific** as to the location of all files and media and all protections that will be in place.

☐ Completed and signed Data Use Agreement (DUA)

☐ Documentation of IRB approval

☐ Completed and signed Request form for restricted variables (if applicable)

☐ Letter from funder (if applicable)

Please send any questions to the SEER-Medicare contact at yaniskoe@imsweb.com

APPLICATION FOR SEER-MEDICARE DATA
(Please complete all information in this form)

I. Contact information

Project Title: < enter project title >

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

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

Alternate contact: Student / fellow contact / assistant / Co-PI: (indicate type)

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

II. Project Description:

A. Title

B. Brief overview of your project (one or two sentences)

C. Cancer sites being requested (e.g.Lung):

D. Description of the Project (between 1-5 pages). This description **must** include:
- statement of main hypothesis/research question
- description of study subjects and cancer sites/phases to be included in the analysis
- brief explanation of how key components of the study will be obtained/identified within the PEDSF and/or claims data– specifically:
  - cohort selection criteria
  - covariates
  - outcomes
- a list of requested files and how each will be used, (for example: MEDPAR will be used …. NCH will be used….)
- how the 5% population (non-cancer and/or other cancer) will be used, if requested
- description of the personnel involved
- timeline for completion
- references can be included, if relevant.

E. Data Storage and Protection:
The **preferred method** of data storage is on an institutional server with all the protections that provides. If you are choosing an alternate data storage method, please provide the rationale as to why you made that choice. Please be aware that Cloud Storage of data does NOT meet privacy rules and will not be approved for storing SEER-Medicare data.

This section must include the following items:
- the **specific location** of the data and where/how the data will be stored
- details on how the data will be protected from unauthorized access.
- information on the storage/protection of the media you receive containing the original files.
- assurances that no attempt will be made to identify individual patients, hospitals or physicians
- assurances that publications and presentations of the data will not allow identification of patients, hospitals or physicians.

F. Funding Source: If your organization is a consulting firm, contractor, or pharmaceutical company, then your application must include a letter from the funder indicating that you are free to work and publish your findings without limitations by the funder. This letter must come from a person in authority on company letterhead.

G. Restricted Variables: Selected variables are not released without the permission of the Principal Investigator of each of the SEER Registries. These variables include census tract of the patient, zip code of the patient, physician or hospital, and unencrypted provider numbers. If you are requesting access to any of these variables, you must include the justification in your description of the project and also submit the completed request form for restricted variables. Please see http://healthservices.cancer.gov/seermedicare/privacy/variables.html for information. NCI will provide a researcher with contact information for each of these registries; however it is the responsibility of the researcher to obtain permission from each registry.

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

Name of file	Years available	Years requested
Patient Entitlement and Diagnosis Summary File (PEDSF) (SEER cases)	1973-2013	
Summarized Denominator File (SUMDENOM) 5% non-cancer sample	1991-2013	
5% PEDSF ** 5% cancer sample (Medicare enrollment data only)	1991-2013	
MEDPAR	1991-2014	
NCH - Carrier (physician/supplier)	1991-2014	
Outpatient	1991-2014	
Home Health (HHA)	1991-2014	
Hospice	1991-2014	
Durable Medical Equipment (DME)	1994-2014	
Part D Event (PDE)	2007-2014	
Chronic Conditions Flag	1999-2014	
Hospital File	1996, 1998, 2000-2014	

(****** To receive the 5%PEDSF, you must also be approved to receive the SUMDENOM file. In this case, your SUMDENOM file will include persons in the 5% sample who never had any cancer (SUMSTAT = 1) and patients in the 5% sample reported to have cancer, but no additional information (SUMSTAT = 2).

Note: Medicare claims prior to 1998 are available only for cases diagnosed with cancer before 2003. Cases diagnosed 2003-2005 have claims from 1998+; cases diagnosed 2006 -2007 have claims from 2000+; cases diagnosed 2008-2009 have claims from 2002+; cases diagnosed 2010-2011have claims 2004+; Cases diagnosed 2012-2013 have claims 2006+

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 number of locations where the data are located shall also be minimized and specific location details 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 http://csrc.nist.gov/drivers/documents/appendix_iii.pdf), 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 does not meet privacy rules and is not acceptable for storing SEER-Medicare data.

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 institution's server.

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 move to a different institution, you must contact NCI in writing prior to moving for instructions on how to handle the SEER-Medicare data. You may not duplicate any SEER-Medicare files prior to moving nor can you take SEER-Medicare data with you without written permission from NCI. If you chose not to take the data with you, you must destroy the files or designate a new PI prior to moving.

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 the applicable SEER registries.

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. In the event that 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. 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. 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.

13. You agree to provide NCI with a copy of all manuscripts to be submitted for publication prior to submission. You further agree not to submit such findings to any third party until receiving NCI's approval to do so. NCI agrees to make a determination about approval and to notify you 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. NCI may withhold approval for publication only if it 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.

14. 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 return the data and all derivative files 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.

15. 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.

Please indicate the SEER-Medicare files you will use:

☐	Patient Entitlement and Diagnosis Summary File) (PEDSF)	Years	
☐	Summarized Denominator File (SUMDENOM)	Years	
☐	MEDPAR	Years	
☐	NCH - Carrier (physician/supplier)	Years	
☐	Outpatient	Years	
☐	Home Health Agency	Years	
☐	Hospice	Years	
☐	Durable medical equipment	Years	
☐	Part D Event (PDE)	Years	
☐	Chronic Conditions Flags	Years	

These files will include:

☐ Cancer cases ☐ Non-cancer cases

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

Request for restricted or unencrypted variables on the SEER-Medicare file

Investigator Name:	
Organization:	
Telephone:	
Email	
Project title:	

Project Abstract:

Check the variable(s) requested. Please be aware that:

- Unencrypted Census tracts and ZIP Codes are NOT needed to link the Census data to the PEDSF or SUMDENOM information. All Census tracts and ZIP Codes on SEER-Medicare data are encrypted in a uniform manner so Census data can be merged using the encrypted variables. Selected Census data from 1990, 2000 and ACS 2008-2012 by ZIP code and Census tract are provided with every data request.

- Unencrypted hospital numbers are NOT needed for volume outcomes studies. All provider numbers on SEER-Medicare data are encrypted in a uniform manner so volume can be calculated from the encrypted numbers.

☐	patient ZIP Code	☐	patient Census tract
☐	Provider/ Hospital ZIP Code	☐	unencrypted hospital provider numbers (NPIs)

Please describe how the requested variables will be used in your project (ex. unencrypted patient and hospital Zip code will be used to determine distance to care)

This project is funded by:	☐ NCI	☐ DOD	☐ Other (please specify):

☐ I agree that if these variables are released to me that they will not be used to identify any individual cancer patient, hospital or physician. I will publish findings from this analysis at a sufficient level of aggregation to make it impossible to identify individual patients and providers, and I will not make public any information that may result in the identification by others of individual patients, hospitals or physicians. I understand that I can only access the SEER-Medicare data to work on the project as described in my application. Furthermore, the dataset with restricted variables can be used only for this particular project and cannot be used for any subsequent analysis.

Date	Investigator's printed name	Investigator's signature

日本版 SEER 研究班平成 29 年度第 1 回班会議資料

金沢医科大学医学部公衆衛生学
西野　善一

担当課題：「がん診療情報」の収集の仕組みの提案に関し、SEER のために収集されたデータを第三者の研究者に提供可能な理由、特に病院等での情報取得時の本人通知及び本人同意の有無

- NCI から第三者の研究者への SEER データ提供と「目的外利用の制限」、「第三者提供の制限」との関連

  - 各登録から NCI への SEER データ提供について本人同意または本人への通知があるか？
  - NCI から第三者の研究者への SEER データ提供について本人同意または本人への通知があるか？
    →アメリカの各州では登録にあたっての本人同意は不要。一部の州では州法でがん登録室から患者への登録通知義務（Oregon）、医療機関から患者へのがん登録についての説明義務（California, Virginia, Washington）有[1, 2]

  - 各登録から NCI への SEER データ提供が本人の同意がなくても可能な法的根拠は何か？
  - NCI から第三者の研究者への SEER データ提供が本人の同意がなくても可能な法的根拠は何か？
  - 個人の秘匿性との関連において NCI から第三者の研究者への SEER データ提供が可能な範囲は？
    - ✓ HIPAA (Health Insurance Portability and Accountability Act) Privacy Rule
      - PHI (protected health information)
      - de-identified data set
      - limited data set

  - NCI はどのような手続きで各登録から SEER データを入手しているか？

[1] 永本裕子. アメリカにおける地域がん登録（前半）. 厚生労働科学研究費補助金第３次対がん総合戦略研究事業「地域がん登録の法的倫理的環境整備に関する研究」平成 17 年度総括・分担研究報告書. 10-25, 2006.
[2] 森本直子. アメリカにおける地域がん登録（後半）. 厚生労働科学研究費補助金第３次対がん総合戦略研究事業「地域がん登録の法的倫理的環境整備に関する研究」平成 17 年度総括・分担研究報告書. 26-42, 2006.

NATIONAL CANCER INSTITUTE
Division of Cancer Control & Population Sciences

Healthcare Delivery Research Program

SEER-Medicare Linked Database

Announcements

- The 2016 SEER-Medicare linkage is now available. Please be sure to use the current version of all forms.
- NCI has launched the Healthcare Delivery Research Program (HDRP), which oversees SEER-Medicare, to address new challenges and opportunities for cancer control research in the context of health care systems.
- The 2016 SEER-Medicare linkage includes PSA data for cases diagnosed in 2010-2013. SEER-Medicare Data Users are advised not to use PSA data that were previously released.

Related Datasets

- New SEER-MHOS flag in SEER*Stat!
- SEER-MHOS data for cohort 15 are now available!

About the SEER-Medicare Database

- Brief Description of the Database
- How the SEER & Medicare Data are Linked
- Publications Using SEER-Medicare Data

The SEER-Medicare Data Files

- About the Data Files
- SEER Program & Data
- SEER-Medicare Policies
- Medicare Enrollment & Claims Data
  - Medicare Enrollment Data
  - Medicare Claims Files
- Summary Table of Available Medicare Data
- Provider Files
- File Sizes & Distribution
- Number of Cases for Selected Cancers
- HCPCS Tables
- Frequency of Prescription Drugs on Durable Medical Equipment Files
- Number of Part D Enrollees
- Frequency of Prescription Drugs on Part D Event Files
- Number of Cancer Patients with Diagnosis Codes

Obtaining the SEER-Medicare Data

- Overview of the Process for Obtaining the Data
- Instructions & Data Use Agreements
- Instructions for Requesting New Data / Changes for Previously Approved Projects
- Proposal Review Process
- Requirements Following Receipt of Data
- Cost of Acquiring SEER-Medicare Data
- SEER-Medicare Data Cost Calculator

Analytic Support for Researchers

- Measurement & Methods
- Identification of Diagnosis & Procedure Codes
- Procedure Codes for SEER-Medicare Analyses
- Comorbidity Index Overview
- Cancer Testing Covered by Medicare
- Defining the Date of Diagnosis & Treatment
- Geographic Location of Care
- Method to Calculate Hormone Therapy for Men with Prostate Cancer
- SEER-Medicare Data Limitations
- SEER-Medicare Training
- Resources for More Assistance
- Measures that are Limited or not Available in the Data
- NPI-UPIN Physician Crosswalk File

Privacy & Confidentiality Issues

- Encrypted Variables
- IRB Approval & HIPAA Regulations
- Laptops & Other Portable Media

Contact SEER-Medicare

- Contact SEER-Medicare

Programming Support

- SAS Input Statements
- Using SEER*Stat to Analyze SEER-Medicare Data
- Calculation of Comorbidity Weights
- Important Message about Merging Cancer Data
- Data File Changes in the 2016 Linkage (PDF, 257 KB)

Last Updated: 20 Apr 2017

NATIONAL CANCER INSTITUTE
Division of Cancer Control & Population Sciences

Healthcare Delivery Research Program

IRB Approval & HIPAA Regulations

Researchers who wish to use the SEER-Medicare data are required to obtain IRB approval prior to SEER-Medicare data being released to them. A full IRB review is not required. Many IRBs, including NIH's Office of Human Subjects Research, have determined that the SEER-Medicare data are exempt (45 cfr 46.101(b)(4)).

Researchers who wish to use the SEER-Medicare data may have concerns about complying with the Health Insurance Portability and Accountability Act (HIPAA) regulations. The SEER-Medicare data contain information about geographic location at the county level as well as dates of receiving health care services. Because of these variables, the SEER-Medicare data are considered by HIPAA requirements as a limited data set, which requires that investigators sign a Data Use Agreement prior to receiving the data. This exception allows for the release of the SEER-Medicare data without obtaining authorization from individual patients (see Federal Register, August 14, 2002, pg 53235). However, because the SEER-Medicare data are a limited data set, investigators who have the data may not share these files with other investigators. Investigators who are contacted by colleagues who wish to use their data should ask their colleagues to contact SEER-Medicare.

Last Updated: 24 Feb 2017

NATIONAL CANCER INSTITUTE
Division of Cancer Control & Population Sciences

Healthcare Delivery Research Program

SEER-Medicare: Encrypted Variables

Physician & Hospital Identifiers

The Center for Medicare and Medicaid Services (CMS) and the SEER registries require that the identity of physicians and other health care providers be protected. The SEER registries require that the identity of hospitals also be protected. Therefore the physician and hospital variables on the SEER-Medicare claims are encrypted. This includes the Unique Physician Identification Number (UPIN), National Provider Identifier (NPI), the provider Taxpayer ID number (tax_num), and hospital provider number (Provider). These numbers are encrypted in a similar manner across files and years making it possible to track the same hospital or physician in the SEER-Medicare data over time.

Investigators may want information about providers that requires linkage to other data sources by using unencrypted provider numbers. In order for NCI to release unencrypted hospital numbers, investigators must obtain permission from each of the SEER registries as described below. As of July 2014, NCI no longer releases unencrypted UPINs/ NPIs, PINs or Taxpayer Identification Numbers.

Many investigators want to link to data about physicians from the American Medical Association (AMA). NCI has established methods to support such linkages without requiring special permission. In order to link to the AMA data, investigators should complete the following steps:

1. Identify the encrypted provider numbers from the Medicare data. Physicians' identifiers are the UPINs or NPIs found on the carrier and outpatient files.
2. Send the encrypted provider numbers to NCI's information technology contractor, IMS Inc. Please e-mail the provider numbers to Bob Banks. IMS will unencrypt the provider numbers and send them to the AMA.
3. AMA will return to IMS their data linked to the unencrypted provider number.
4. IMS will re-encrypt the file and return to the investigator a file with encrypted provider numbers and the selected AMA variables.

Investigators must negotiate directly with the AMA about the variables needed, terms of use and the cost of any processing. Neither NCI nor IMS are involved in this process.

Researchers who are seeking **AMA data** should **direct any inquires** to AMA's programming contractor, Medical Marketing Services, Inc.:

Tom Lorge
Medical Marketing Services, Inc.
185 Hansen Court, Suite 110
Wood Dale, IL 60191
Phone: 630-477-1564
Fax: 630-350-1896
t-lorge@mmslists.com

Geographic Identifiers

The patient's county of residence is available on the PEDSF (FIP codes) and in the Medicare files (SSA codes). To protect patient and provider identification, NCI encrypts other geographic variables including patient's census tract and ZIP code, physician ZIP code, and hospital ZIP code. Separate files that contain geographically-based (ZIP code and census tract level) socioeconomic information from the 1990 and 2000 Censuses and the 2008 – 2012 American Community Survey are provided and can be matched by the encrypted patient census tract and ZIP code to the claims files. Unencrypted ZIP codes and census tracts can only be released if the investigator obtains permission of each SEER cancer registry.

Exceptions to Allow Release of Unencrypted Variables

If investigators determine that encrypted variables are needed in an unencrypted format for their analysis, they must go through a special approval process. Investigators must submit their completed application form (DOCX, 29 KB) to the SEER-Medicare contact with a detailed justification for access to the unencrypted variable(s). A completed and signed request form (DOCX, 19 KB) and a list of people that will have access to these data must be included with the request. An NCI staff member will review the application. Once NCI supports the request for these variables, investigators must obtain permission from each of the registries prior to release of unencrypted variables for that registry. The SEER-Medicare contact will provide investigators with contact information for the SEER registries. Investigators who are requesting unencrypted variables are encouraged to allow sufficient time to obtain the approval from each SEER registry.

Note: Files with unencrypted variables cannot be stored with regular SEER-Medicare data. In order to combine multiple requests when purchasing data, all requests must have the same permissions for access to any unencrypted variable.

Last Updated: 24 Feb 2017