

II. 分担研究報告

1. 全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究 - 記述疫学研究担当 -
研究分担者 堀 芽久美、片野田 耕太
付録1 SEER Data 利用手続き
付録2 SEER Linked Database 利用手続き
2. 全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究 - 宮城県をモデルとした基盤整備と方策に関する検討 -
研究分担者 金村 政輝
付録1 ユタ州がん登録視察結果
付録2 米国の仕組み
3. 全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究 - 仕組み検討（モデル地域候補） -
研究分担者 大木 いずみ
4. 全国がん登録情報を基盤とした長期記述疫学研究用の特定匿名化情報の整備について - 米国 SEER にデータを提供しているユタがん登録訪問により見えてくる課題 -
研究分担者 伊藤 秀美
5. 全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究 - 記述疫学研究担当 -
研究分担者 林 櫻松
6. 全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究 - 記述疫学研究（国際比較）疫学研究として活用するためのデータベースの条件 -
研究分担者 井上 真奈美
7. 全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究 - 全国がん登録情報利用者に求められる安全管理措置 -
研究分担者 西野 善一

厚生労働科学研究費補助金（がん対策推進総合研究事業）
（分担）研究報告書

全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究

- 記述疫学研究担当 -

研究分担者 堀 芽久美 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 研究員

研究分担者 片野田 耕太 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 部長

研究要旨 SEER 研究データの取得方法、提供データの粒度調査および SEER データの記述統計解析によって SEER 研究データの利用可能範囲、研究活用への有用性を評価した。また、SEER を訪問し、SEER 研究データ提供体制の詳細を調査することで、日本で実現可能なデータ提供体制のあり方を検討した。SEER は研究へのデータ利活用の推進を重視しており、容易な手続きでのデータ取得が可能で、1 症例の罹患例も入手可能となる情報粒度を持つ。組織型別罹患率、稀少がんの罹患率も SEER 研究データから算出可能であり、SEER から入手できる情報は現在日本で必要とされるがん統計の整備に有用であった。日本ではデータ提供に関わる個人情報の保護は大きな課題である。SEER のようなデータ利用手続きの簡易性、情報粒度を実現するには患者、医療者、行政、企業、研究者等が協力してさらなる検討を進めていく必要がある。全国がん登録データの利活用の促進に向けては、一般ユーザーでも理解しやすく使いやすい全国がん登録データ解析用ソフトウェア等の開発が望まれる。

A. 研究目的

Surveillance, Epidemiology, and End Results (SEER) は米国 National Cancer Institute (NCI) の Surveillance Research Program (SRP) が支援するがん対策に資するがん統計情報の提供を目的としたプログラムである。

本研究では SEER で提供されるデータの取得方法、提供データの粒度を調査し、さらに SEER から取得したデータを利用した記述研究を行うことによって、SEER 研究データの利用可能範囲、研究活用への有用性を評価した。また、SEER のデータ提供に対する指針を調査し、日本で実現可能なデータ提供体制のあり方を検討した。

B. 研究方法

SEER ウェブサイト(<https://seer.cancer.gov/>)から、SEER 1973-2014 research data 取得の利用申請を行った。SEER 研究データの取得方法は SEER*Stat から SEER 研究データベースへアク

セスする方法、SEER のウェブサイトからバイナリー形式か ASCII テキスト形式のデータをダウンロードする方法の 2 通りがある。SEER*Stat とは SEER や他のがんに関するデータベースの解析に有用な統計ソフトウェアで、SEER のウェブサイトから誰でも入手可能である。本研究では SEER*Stat からデータを取得した。取得したデータは卵巣がんの組織型別罹患率の算出に利用し、この解析を通して提供データの精度を調査した。

SEER のデータ提供体制の詳細を調査するため、SEER の支援組織である米国 NCI を訪問し、SEER 研究データの管理に関わるスタッフにインタビューを行った。

（倫理面への配慮）

本研究で用いた SEER 研究データは個人情報を含まないため、倫理面への問題は生じないと判断される。

C. 研究結果

SEER*Stat から SEER 研究データベースへアクセスするには、SEER のウェブサイトからオンラインでの利用申請および電子メールでの誓約書の送付が必要であった。オンライン利用申請、誓約書の内容は以下のとおりである。

<利用申請書>

- ・ アクセス方法
- ・ 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
- ・ 申請者情報

<誓約書>

- ・ 研究目的以外で利用しない
- ・ 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
- ・ 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
- ・ 個人を特定しようとするしない
- ・ 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
- ・ SEER の承認を得た人以外にデータを利用させない
- ・ 適切なデータ保管措置を行う
- ・ SEER から提供されるソフトウェアの適切な使用
- ・ 結果公表の際の適切な引用

誓約書の送付後、データ利用が承認されると、SEER*Stat のクライアントサーバーへのログインパスワードがメール送信され、SEER 研究データへのアクセスが可能となる。

SEER*Stat から SEER 1973-2014 research data へアクセスし、卵巣がんの組織型別罹患率を算出した。SEER から取得できるデータでは罹患率は 1 歳刻みの年齢区分が利用可能であるが、人口データは対応していない。年齢区分に関係なく、罹患少数例のセルも表示される。地域分類は County までが利用可能であった。また、County に対応した American Community Survey (ACS) データの利用が可能で、地域別の社会・経済指標

とがん罹患・死亡との関連検討に適用できる。組織型については、ICD-O-3 分類に基づく組織型/性状情報が含まれ、診断年や、他の患者情報（性別、年齢階級、人種等）と組み合わせられた集計が可能であった。

SEER*Stat から取得したデータを用いて算出した年齢階級・人種・組織型別罹患率を図に示す。White, Black, American Indian/Alaska Native(AI/AN), Asian or Pacific Islander(API)別では、40 歳代から 50 歳代の明細胞線癌、膵臓癌において API が他の人種に比較して罹患率が高く、その後の減少傾向が顕著であった。一方、White、Black の罹患率は 50 歳代後半から 60 歳代にピークをむかえるが、その後の減少は API と比較して緩やかである。他の組織型では罹患数が少なく比較や解釈は難しいが、SEER 研究データを活用した場合、組織型別罹患率の算出は容易に行うことができた。SEER データを利用して他国と比較する場合には、組織型登録の不明率などデータの登録の質に関する比較が必要である。

SEER のデータ提供に対する考え方を調査するため、米国 NCI を訪問し、SEER データの管理に関わるスタッフにインタビューを行った。

SEER データ利用手続きの簡易性は、研究への利活用を重視する姿勢により実現していた。SEER データでは少数罹患例も集計可能であるが、たとえ 1 例になっても、その患者を知らない限り個人が特定される可能性は低いと考えられている。どの程度少数になるとデータの提供を控えるのかといった基準はない。また、SEER は SEER*Stat のような、SEER データの解析に有益なソフトウェアの開発・提供も行っていた。統計に詳しくない一般のユーザーでも、簡単に罹患率や生存率が推計できる仕組みであった。さらに、統計的に正しい解釈ができるようにツールの利用が工夫されている。

D. 考察

SEER 研究データはオンライン申請および誓約

書の提出のみで用意に取得が可能であった。提供されるデータも日本では一般的には入手できない情報粒度を持つ。SEER 研究データの情報粒度は今後わが国で整備が必要とされる稀少がんの罹患率や組織型別罹患率の算出に有用なものであった。ただし、SEER は米国全体をカバーしたデータではなく、SEER 1973-2014 research data は 18 州のデータをまとめたものである。これに対して、日本で実施されている全国がん登録は日本全体をカバーしており、データの悉皆性は現在の SEER にない強みである。

日本においてデータ提供に関わる個人情報の保護は大きな課題である。SEER の研究への活用を重視する姿勢は日本より強い。SEER と同程度の情報粒度や利用手続きの簡易性を実現するには、患者、医療者、行政、企業、研究者の各方面からの意見収集・検討が必要であろう。今後、全国がん登録データの提供に対応して SEER*Stat のような誰もが簡単に入手・利用できるソフトウェアの開発はがん登録データの利活用促進に向けて有用であると考えられる。

E. 結論

SEER 研究データは簡易な手続きで粒度の高い

情報を取得することができる。SEER が提供する情報粒度は、今後わが国での整備が期待されるがん統計情報の整備に有益であった。しかし、SEER と同程度の情報粒度や手続きの簡易性を実現するには更なる検討を要する。

今後、全国がん登録データの提供に対応して利活用促進に向けたソフトウェアの開発の実現が望まれる。

F. 健康危険情報

総括研究報告書にまとめて記載する

G. 研究発表

1. 論文発表
なし
2. 学会発表
なし

H. 知的財産権の出願・登録状況

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

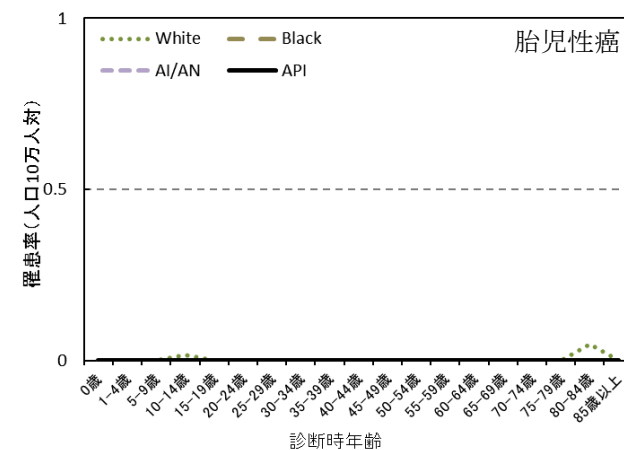
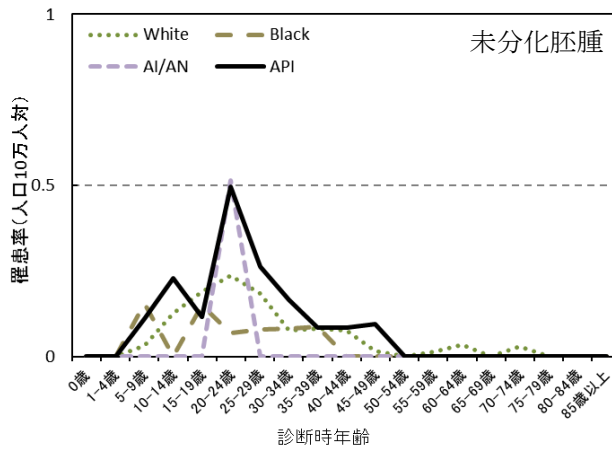
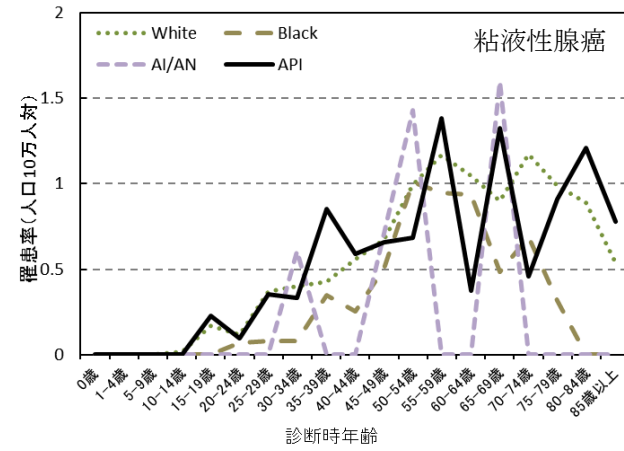
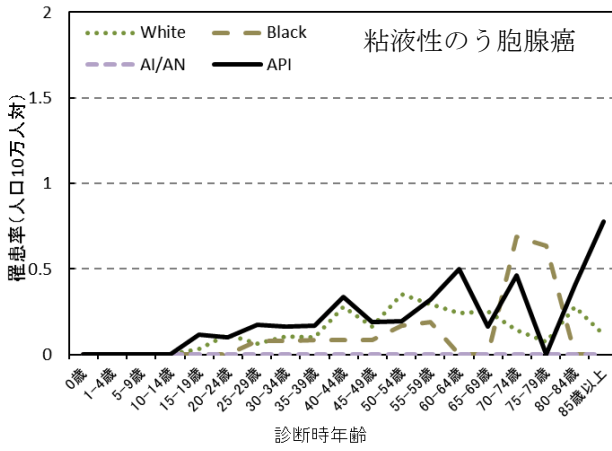
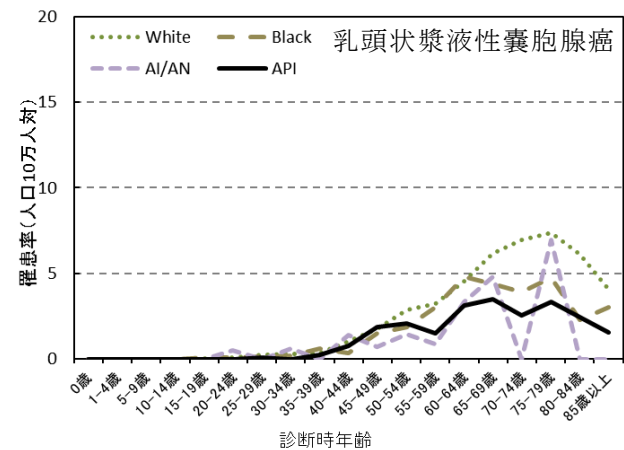
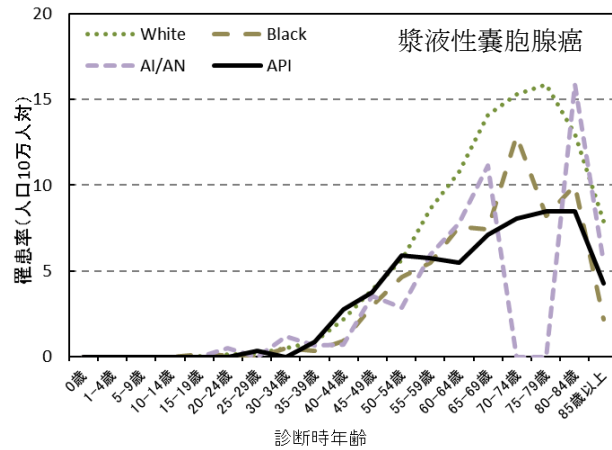
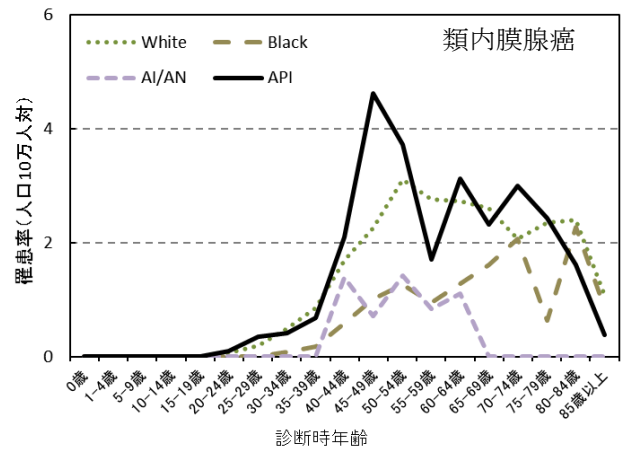
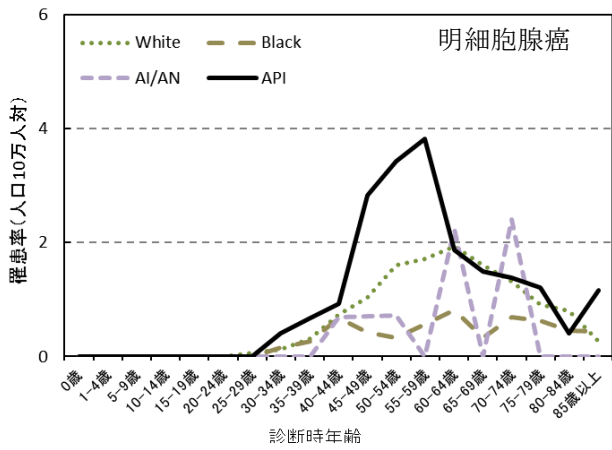


図 卵巣がんの組織型別・人種別・年齢階級別罹患罹患率

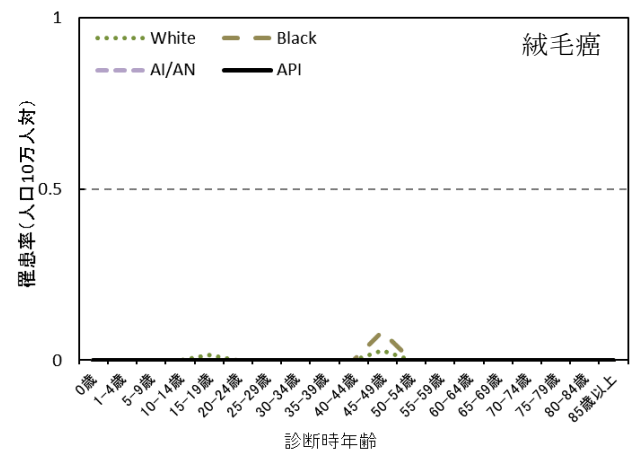
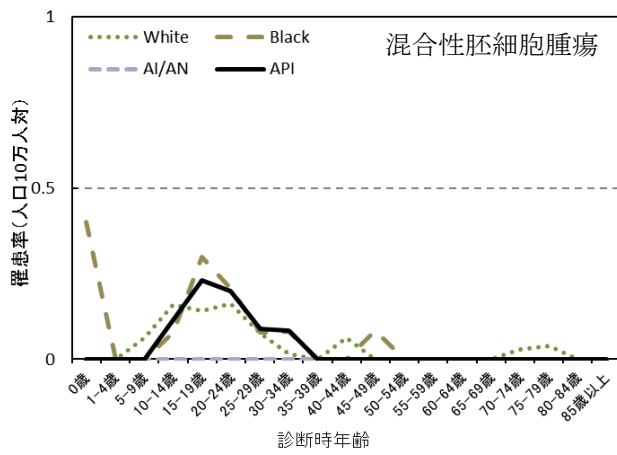
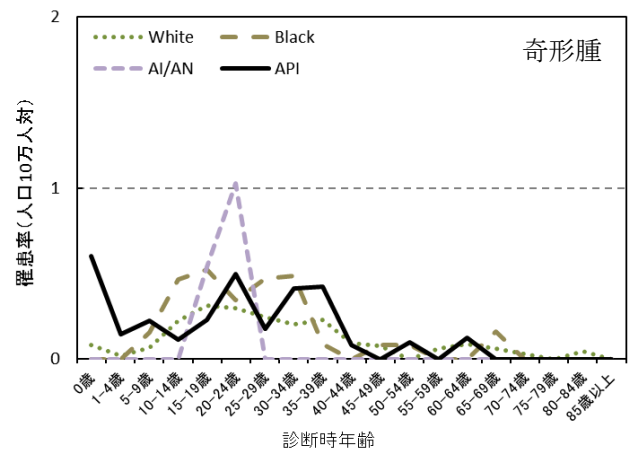
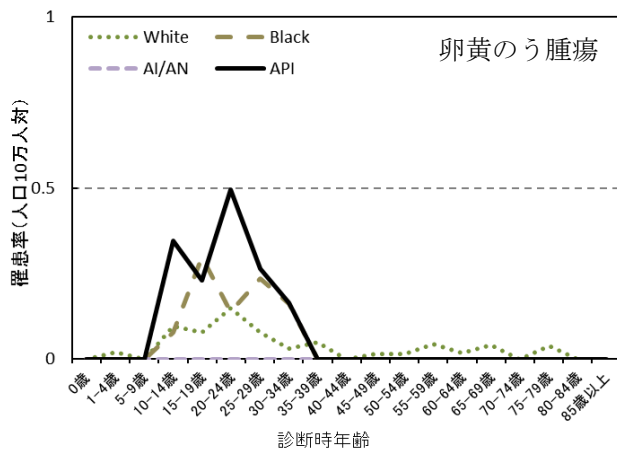
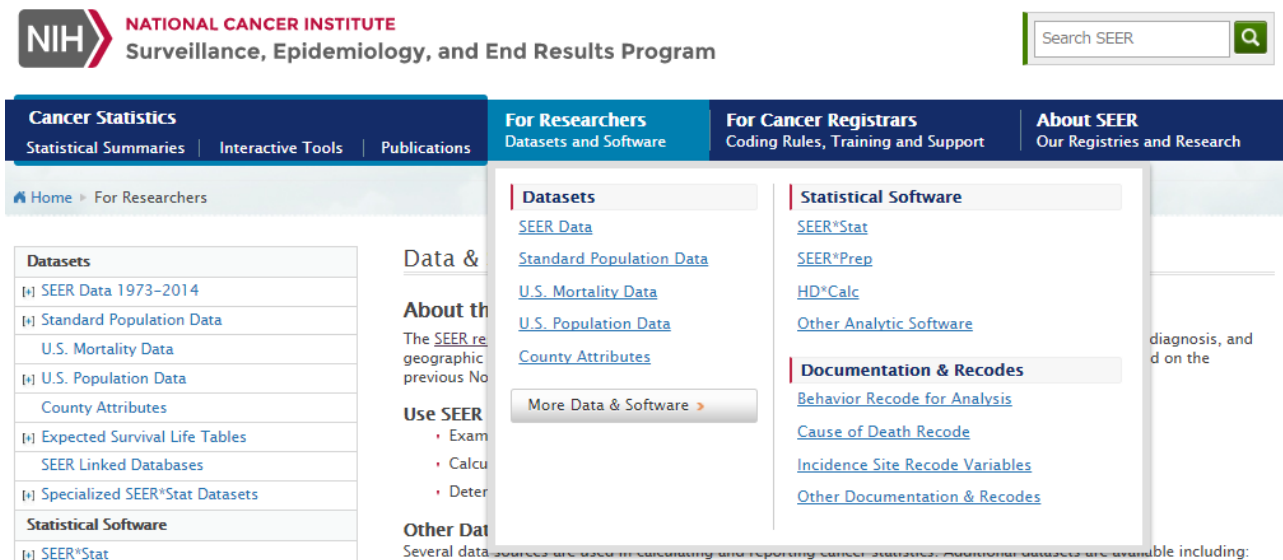


図 卵巣がんの組織型別・人種別・年齢階級別罹患率 (つづき)

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するしない
 - 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

☒ 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.

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

☒ 2

SEER 1973–2014 Research Data Access Request

First name: *

Middle initial:

Last name: *

Organization:

Address: *

City: *

State: *

Zip: *

Country: *

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 specify:

What is your purpose for using the data?:

*** required item**

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	0
8000/1: Neoplasm, uncertain whether benign or malignant	0	0	0	0	0	0	0
8000/2: Neoplasm, in situ	0	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	0
8001/1: Tumor cells, uncertain whether benign or malignant	0	0	0	0	0	0	0
8001/3: Tumor cells, malignant	0	0	1	0	0	0	0
8002/3: Malignant tumor, small cell type	0	0	0	0	0	0	0
8003/3: Malignant tumor, giant cell type	0	0	0	0	0	0	0
8004/3: Malignant tumor, spindle cell type	0	0	0	2	0	0	0
8005/3: Malignant tumor, clear cell type	0	0	0	1	0	0	0
8010/0: Epithelial tumor, benign	0	0	0	0	0	0	0
8010/2: Carcinoma in situ, NOS	0	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	0
8011/3: Epithelioma, malignant	0	0	0	0	0	0	0
8012/2: Large cell carcinoma in situ	0	0	0	0	0	0	0
8012/3: Large cell carcinoma, NOS	0	0	0	0	0	0	0
8013/3: Large cell neuroendocrine carcinoma	0	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.

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 2010-2011 have 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:

<input type="checkbox"/>	Patient Entitlement and Diagnosis Summary File) (PEDSF)	Years	
<input type="checkbox"/>	Summarized Denominator File (SUMDENOM)	Years	
<input type="checkbox"/>	MEDPAR	Years	
<input type="checkbox"/>	NCH - Carrier (physician/supplier)	Years	
<input type="checkbox"/>	Outpatient	Years	
<input type="checkbox"/>	Home Health Agency	Years	
<input type="checkbox"/>	Hospice	Years	
<input type="checkbox"/>	Durable medical equipment	Years	
<input type="checkbox"/>	Part D Event (PDE)	Years	
<input type="checkbox"/>	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.

<input type="checkbox"/>	patient ZIP Code	<input type="checkbox"/>	patient Census tract
<input type="checkbox"/>	Provider/ Hospital ZIP Code	<input type="checkbox"/>	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:	<input type="checkbox"/> NCI	<input type="checkbox"/> DOD	<input type="checkbox"/> 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

厚生労働科学研究費補助金（がん対策推進総合研究事業）
（分担）研究報告書

全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究
－宮城県をモデルとした基盤整備と方策に関する検討－

研究分担者 金村 政輝 宮城県立がんセンター研究所 がん疫学・予防研究部

研究要旨

本研究班では、米国の **SEER** のように全国がん登録を基盤とした長期にわたる疫学研究を実現するためのモデルの提案を目指している。今年度は、米国ユタ州がん登録を対象に情報収集を行い、現地調査を行った。ユタ州では、**SEER** などのプログラムからの資金提供を受け、スタッフの雇用が可能となっており、関係団体による支援によってデータの内容や形式の定義、トレーニングの機会の提供と認証が行われ、質が担保されるとともに、標準化されたデータの登録・提供が可能となっていた。州のがん登録は、各州の法律・規則に基づき運営されているが、ユタ州では、様々な工夫をしながら情報の収集を行っており、州の保健当局との取り決めに従い、**Advisory Research Committee** による助言を受けながら、データの公表・提供が行われていた。我が国において **SEER** のような仕組みを実現するための課題としては 11 項目が抽出された。宮城県をモデルとして、全国がん登録の届出対象項目にはないがん診療情報を追加で収集するためには、大きく 2 つの方法が考えられた。特定匿名化情報の提供においては、仕組み自体は全国共通のものとして整える必要があるものの、少数例の取扱いが宮城県においても悩ましい課題であり、実際にモデル地域での取り組みなどを通して解決することが必要と考えられた。

A. 研究目的

米国では、全米を対象としたプログラムである **Surveillance, epidemiology, and End Results (SEER)** が長期にわたり実施されており、この **SEER** によるデータを用いたがん疫学研究の研究成果が数多く報告されている。全国がん登録が開始された我が国において、今後、この **SEER** のような全国がん登録を基盤とした長期にわたる疫学研究の実現が期待されている。そのためには、診療に関するより詳細な情報を収集し、提供する仕組み（がん診療情報の追加収集・提供）を構築し、匿名化した上で情

報提供すること（特定匿名化情報の提供）が必要である。

本研究班では、長期にわたり地域がん登録を実施し、高い精度での情報収集が行われてきた愛知、栃木及び宮城の 3 県をモデル地域と位置付け、モデルの提案を目指している。本年度は、宮城県でのモデル事業の実施を念頭に置いた上で、その実現のための課題抽出を行い、今後必要な事項についての提案を行うため、以下の研究活動を行った。

B. 研究方法

(1) ホームページ等調査・米国視察

米国の SEER における情報収集と提供の仕組み、特に、州のがん登録室と医療機関との間において、詳細ながん登録情報、診療情報の収集をどのように実現しているのかをホームページ等から調査を行った。その後、米国ユタ州のがん登録室を実際に訪問し、情報収集と提供の仕組みについて現地調査を行った。現地調査の結果については、モデル地域として参加している他の仕組み検討班のメンバー（※）と分担し、班会議に報告を行うとともに、記述疫学検討班などの共同研究者と情報交換・意見交換を行った。

(2) 課題抽出と提案

上記の結果を踏まえ、宮城県でのモデル事業の実施を念頭に置いた上での課題抽出を行うとともに、今後必要な事項についての提案を行った。

C. 研究結果

(1) ホームページ等調査・米国視察

ホームページ等の調査を行い、現地調査を行った結果を【別紙 1】に示す（ただし金村の担当分のみを提示）。米国においては、SEER によるプログラムとは別に、CDC が資金提供している NPCR（National Program of Cancer Registries）というプログラムがあり、前者は 15 州、後者は 46 州と参加規模の違いがある。疫学研究では、SEER のデータが使われていることが多いが、開始から長期にわたること、最終的な目標が survival であることなどの違いがあることが確認出来た。いずれのプログラムからも資金提供が行われており、登録士 Registrar の雇用が可能となっていた。関係団体による支援によって、データの内容や

形式の定義、トレーニングの機会の提供と認証が行われることで、質が担保され、標準化されたデータの登録・提供が可能となっていた。米国では、各州独自の法律・規則によってがん登録について種々の規定がなされ、それに基づき各州のがん登録が運営されているが、ユタ州では、実務上、様々な工夫をしながら情報の収集を行うとともに、ユタ州の保健当局との取り決めに従い、Advisory Research Committee による助言を受けながら、データの公表・提供を行っていた。

(2) 課題抽出と提案

上記の結果については、仕組み検討班の小班会議において情報交換・意見交換を行った後、課題抽出を行い、視察結果と合わせて班会議に報告した（【別紙 2】。ただし、金村の担当分のみを提示）。課題としては全部で 11 項目が抽出された（①外部資金によるスタッフの継続雇用、②収集部門と情報提供部門それぞれへの専門職種マネージャーの配置、③対象症例抽出用のツールの開発、④病理報告の様式の標準化、⑤病理報告の電子データ化と収集の仕組みの構築、⑥非匿名化情報の研究利用を可能とする制度設計、⑦二次利用の方法の仕組みづくり、⑧全国がん登録届出対象外項目抽出のためのデータ収集機関・会社等の利用、⑨データ収集機関・会社等へ支払う予算・財源の確保、⑩届出対象外項目の収集の際のデータの標準化の担保、⑪病院に対する情報提供やトレーニングの機会の提供）。その後、班会議での議論を踏まえ、宮城県でのモデル事業の実施を念頭に置いた上で、その実現のための課題抽出と具体的な提案を行った（【別紙 3】）。全国がん登録の届出対象項目にはないがん診療情報を追加で収集する

ためには、大きく 2 つの方法が考えられた。ひとつは既存の院内がん登録全国集計を活用する方法、もうひとつは、都道府県が主体となって実施する方法で、後者の場合には、①必要性、②周囲からの期待や要請、③都道府県の考え方・熱意・環境、④地域に対する後押しにより、その実現性が変わるものの、合意形成を図りながらの実施が必要であり、サンプリング調査では、漏れがないように、さらに、出張採録も検討することが必要と考えられた。特定匿名化情報の提供においては、仕組み自体は全国共通のものとして整える必要があるものの、少数例の取扱いが宮城県においても悩ましい課題であり、実際にモデル地域での取り組みなどを通して解決することが必要と考えられた。

D. 考察

米国視察を通して課題抽出を行ったが、彼我の違いは大きく、解決すべき課題として 11 項目が抽出された。これらは、スタッフの雇用と配置 (①、②)、抽出・収集のための支援・標準化・機関等の活用 (③～⑤、⑧、⑩、⑪)、そのための仕組みづくりや予算・財源の確保 (①、⑥、⑦、⑨) の 3 つに区分される。このうち、予算・財源なしに着手できる部分については、今後、その実現に向けて具体的な提言につなげていく必要がある。しかし、予算・財源を伴うものについては、直ちに実現することが難しく、今後の課題である。

全国がん登録の届出対象項目にはないがん診療情報を追加で収集するためには、大きく 2 つの方法が考えられたが、実際に宮城県で実施する際には、合意形成などの課題があり、また、特定匿名化情報の提供に

おいては、少数例の取扱いが課題であることが確認された。今後、宮城県を対象にモデル事業を実施する際には、米国視察を通して抽出した課題と併せて解決する必要があると考えられた。

E. 結論

米国の SEER のように全国がん登録を基盤とした長期にわたる疫学研究を実現するため、米国ユタ州を対象として、情報収集を行い、現地調査を行った。我が国において SEER のような仕組みを実現するための課題として 11 項目が抽出された。宮城を対象にモデル事業として届出対象項目にはないがん診療情報を追加で収集するためには、大きく 2 つの方法が考えられた。一方で、特定匿名化情報の提供においては、少数例の取扱いが地域としても悩ましい課題であり、実際にモデル地域での取り組みなどを通しての解決が必要と考えられた。

※ 仕組み検討班 (50 音順)

伊藤 秀美 (愛知県がんセンター研究所、モデル地域候補)

大木いずみ (栃木県立がんセンター、モデル地域候補)

柴田亜希子 (国立がん研究センター、研究代表者)

松田 智大 (国立がん研究センター)

G. 研究発表

1. 論文発表

- 1) Liu Y, Shu XO, Wen W, Saito E, Rahman MS, Tsugane S, Tamakoshi A, Xiang YB, Yuan JM, Gao YT, Tsuji I, Kanemura S, Nagata C, Shin MH, Pan WH, Koh WP, Sawada N, Cai H,

Li HL, Tomata Y, Sugawara Y, Wada K, Ahn YO, Yoo KY, Ashan H, Chia KS, Boffetta P, Inoue M, Kang D, Potter JD, Zheng W: Association of leisure-time physical activity with total and cause-specific mortality: a pooled analysis of nearly a half million adults in the Asia Cohort Consortium. *Int J Epidemiol*.doi:10.1093/ije/dyy024.2018

なし

- 2) 金村政輝: がん転移学(上)Ⅱがんの疫学
2. 臓器別のがん転移に関する疫学情報
(3) 乳がん. 日本臨床 75 巻増刊
841-45.2017

2. 学会発表

- 1) 長澤あかね, 佐藤優希, 目崎はる香, 齋藤美登里, 佐々木真理子, 鈴木智子, 鈴木美裕紀, 只野尚子, 金村政輝: 宮城県における全国がん登録への対応ー届出病院支援の取り組みー (第 2 報) . 日本がん登録協議会第 26 回学術集会, 松山, 2017 年 6 月
- 2) 金村政輝, 佐藤真弓, 新田彩子: 全国のがん診療連携拠点病院における院内がん登録の実施体制. 日本がん登録協議会第 26 回学術集会, 松山, 2017 年 6 月
- 3) 佐藤真弓, 新田彩子, 金村政輝: 自動集計プログラムの構築と全国集計公表データを使った他院の比較. 日本がん登録協議会第 26 回学術集会, 松山, 2017 年 6 月
- 4) 金村政輝: 全国のがん診療連携拠点病院における院内がん登録の現状分析ー実施体制と実績との関連. 第 76 回日本公衆衛生学会総会, 鹿児島, 2017 年 10 月

H. 知的財産権の出願・登録状況

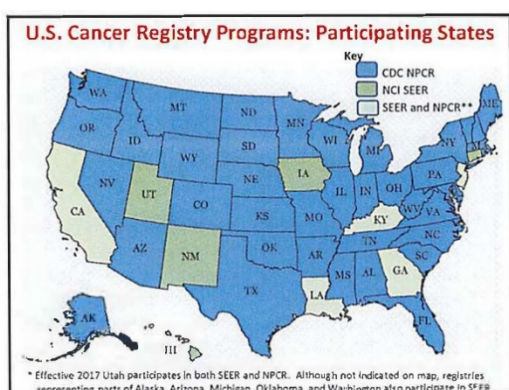
米国における情報収集と提供の仕組み（ユタ州がん登録視察結果）

1 Oversight and Relationship to US Federal and State Agencies

(1) 米国には2つのがん登録プログラムがある

官庁	National Cancer Institute (NCI)	Centers for Disease Control and Prevention (CDC)
がん登録プログラム	Surveillance, Epidemiology, and End Results (SEER)	National Program of Cancer Registries (NPCR)
略称	NCI SEER	CDC NPCR
参加州数	15 ※1970年に9州からスタートし、ユタ州は1973年から参加 ※地域的に偏りあり	46 ※今年6月からユタ州が参加し46州になった
姿勢	・罹患と生存のトレンドを追跡 ※最終的な目標は survival ・がん研究へのデータ提供	・罹患と生存のトレンドを追跡 ・がん対策へのデータ提供
URL	https://seer.cancer.gov/	https://www.cdc.gov/cancer/npcr/index.htm

(2) がん登録プログラムへの参加州



- CDC NPCRに参加
…多くの州に参加
- NCI SEERに参加
…数州に限られる
※ユタ州は6月からNPCRに参加したので図に誤りあり
- SEER及びNPCRに参加
…数州に限られる

(3) SEERとNPCRとそれぞれの州のがん登録の関係

- SEERはNCIとの契約、NPCRはCDCとの協力協定に基づき、参加したがん登録に資金提供を行っている。
※ この資金提供によって、Registrarsの雇用が可能となっている。
- それぞれのがん登録は、SEER及び/又はNPCRに個人同定できない形式でデー

タを提出する。

- SEER と NPCR は、がん登録からのデータを集めて大きなデータセットにし、がん罹患のトレンド評価やがん研究に利用可能としている。

(4) SEER や NPCR での登録方法の発展プロジェクトの役割

- 参加したがん登録は、登録方法の発展と品質保証のための特別プロジェクト(※)に参加申請をする機会が得られる。

※ メールでの情報提供やトレーニングの機会など

- UCR における最近のプロジェクトの例

- SEER Virtual Tissue Repository
- 15-month Resubmission to Improve Completeness
- UCR Data Linkage with Utah All-Payer Claims

※ 米国では日本と保険制度が異なり、保険が変わることもあるため重要

※ データリンクは UCR で行っていて、Social Security Number があるため、リンクは高い確率で一致している。Social Security Number が提供してもらえないときは、リンクには誕生日と名前を利用するが、ラストネームが変わったり、ニックネームや略称などがあるため、困難を伴うことがある。

- Pilot to Collect Screening Information for Breast Cervical Cancer Cases

(5) 登録データの標準化と品質に対する SEER や NPCR の役割

- NCI との契約、CDC との協力協定により、報告対象とするがん、それぞれの症例における項目、データの提出期限が明らかにされている。

- SEER と NPCR は、提出されたデータを完全性(※)と質で評価している。

※ 完全性は過去のデータのレビューの結果により行われている。

(6) 米国における他のがん登録の組織

NCI と CDC という連邦政府の官庁に加え、いくつかの専門家組織がデータの内容と形式の定義、トレーニングと認証についての標準化に関わっている。(※いずれの団体は全く別個のものである。)

- North American Association of Central Cancer Registries (NAACCR)

※ Central registries(population-based registries)が参加し、重要な役割を担っている。

- American College of Surgeons Commission on Cancer (CoC)

※ 外科医の専門団体で、“Cancer Hospital”の認証を行っている。がん登録はこの認証を受けた Cancer Hospital からデータの提供を受けている。

- National Cancer Registrars Association(NCRA)

※ Cancer Registrar が各個人で参加。トレーニングの機会(有料)を提供し、Certified Tumor Registrar (CTR)の認証を行っている。トレーニングはいくつかのコースがあり、期間が異なる。通常、2年間の実務期間の後に認

証の試験を受ける。

※ 州によって違いはあるかもしれないが、CTR の認証を受けることは雇用
に重要となっている。実際に入職する前にもトレーニングを受けている。

□ American Joint Commission on Cancer(AJCC)

※ pathologist が多く、Staging の決定に関与。

(7) がん登録に対する州法の権限

米国では、各州に独自の法律と規則があり、サーベイランス目的での健康データの
報告について規定している。

(8) がん登録に対する州法の権限の特徴

典型的には、以下のような記載になっている。

□ 病院及び他のケア提供施設は、がんの報告が求められている。

□ 中央登録室は、個人同定が可能ながんのデータを受け取る。

※ 複数の施設を受診した同一の患者を 1 人に同定するためには、名前、social
security number、誕生日が必要。

□ 中央登録室は、プライバシーと機密性を保護することが求められている。

(9) Utah Cancer Registry (UCR)

□ ユタ州全域を対象とした population-based のがん登録

□ 1966 年からユタ大学で運営

□ 1973 年から SEER プログラムに参加

□ 2017 年から CDC の NPCR に参加

※ ユタ州は広大で、大部分の人々は限られたエリアである主要な高速道路沿いの都
市部に済み、ヘルスケアへのアクセスもよいが、その他は小さなタウンで、ネイテ
ィブ・アメリカンなども住んでいる。有名なところでは、ナバホ国があり、ナバホ・
インディアンが住んでいる。

※ Ethnicity については、例えば、提供を受けた 2004-2013 の報告書の 7 ページの
table3 に Hispanic or Latino として、罹患数、罹患率、死亡数の記載があるが、ネ
ィティブ・アメリカンの数は少ない。

※ 都市部以外ではヘルスケアへのアクセスは必ずしもよくなく、ネイティブ・アメリ
カンでは複雑な問題だが、すべてがソルトレイクシティなどの都市部の施設を受
診するわけではなく、コロラド州、ラスベガス、ネバダ州などの州外への受診も多
い。

(10) ユタ州におけるがん登録に対する報告についての権限付与

ユタ州では、以下の州法と行政規則がある。

□ 州法：Utah Code Title 26 Utah Health Code, Chapter 5 Chronic Disease Control,
Section3,

System for detecting and monitoring diseases established by department

□ 行政規則：Utah Administrative Code Rule R384-100, Cancer Reporting Rule

<https://rules.utah.gov/publicat/code/r384/r384-100.htm>

(11) 州法 : Utah Code Title 26-5-3

System for Detecting and Monitoring Diseases Established by Department

- ❑ “The department (of health) shall develop and maintain a system for detecting and monitoring chronic diseases within the state and shall investigate and determine the epidemiology of those conditions which contributed to preventable and premature sickness, or both, and to death and disability.”

(12) 行政規則の key language: Reporting

- ❑ “All hospitals, radiation therapy centers, pathology laboratories licensed to provide services in the state, nursing homes, and other facilities and health care providers involved in the diagnosis or treatment of cancer patients shall report or provide information related to a cancer or reportable benign tumor to the Registry.”

(13) 行政規則の key language: Responsibilities

- ❑ “Cancer records are managed by the Utah Cancer Registry (Registry) on behalf of the Utah Department of Health.”
- ❑ “The Utah Department of Health retains ownership and all rights to the records.”
- ❑ “All reports required by this rule are confidential under the provisions of Title 26, Chapter 3 and are not open to inspection except as allowed by Title 26, Chapter 3. The Registry shall maintain all reports according to the provisions of Title 26, Chapter 3.”

(14) UCR の経験から Reporting Rule についてのコメント

- ❑ このルールは、データの形式や変数のような具体的にまでは言及していない。
- ❑ 言い回しが具体的ではないため、SEER、NPCR、NAACCR などが登録における変数やデータ形式に変更を要求してきても、ルールの変更を必要としない。

(15) 州内の病院に対する Reporting Rule の効果

- ❑ UCR への報告のために個々の症例について詳細な情報を収集することを目的として、病院は cancer registrar を雇用する、あるいは、このようなサービスを提供する会社と契約している。病院はこの「データの抽出」の費用を支払わなければならない。
- ❑ ユタ州の病院やヘルスケア提供施設は、全般的に報告に対してとても協力的である。UCR は日常的に病院やヘルスケア提供施設と接触し、タイムリーに報告するよう奨励している。UCR とユタ州はこれまで登録規則の法的な強制力を行使したことはない。

(16) UCR における秘密保持とデータの公表 : ユタ州法から

- ❑ Title 26, Utah Health Code
- ❑ 26-3-7. Disclosure of health data -- Limitations.
<https://le.utah.gov/xcode/Title26/Chapter3/26-3-S7.html>

- 情報公開は以下が必要
 - ・ With consent of the individual or NOK ※NOK: next-of-kin 近親者
 - ・ To a government entity if consistent with purpose for which data were collected,
or
 - ・ “for a specified period, solely for bona fide research and statistical purposes”,

(17) UCR のデータ公表についての概略

Utah Department of Health(UDOH)

- 登録データは、UDOH が所有
- UCR は 1 年分のデータを UDOH に送り、概要報告を提供し、公衆衛生の目的のために UDOH とのデータリンケージを行う。
- UDOH とユタ大学との合意覚書により、他の権限付与されたデータの公表について明確化されている。
- UCR の指導者は月 1 回 UDOH と会合を行っている。

(18) 追加事項

UCR の Advisory Research Committee は、

- 年 2 回会合を行っている。
- UDOH、病院、研究者からの代表が参加している。

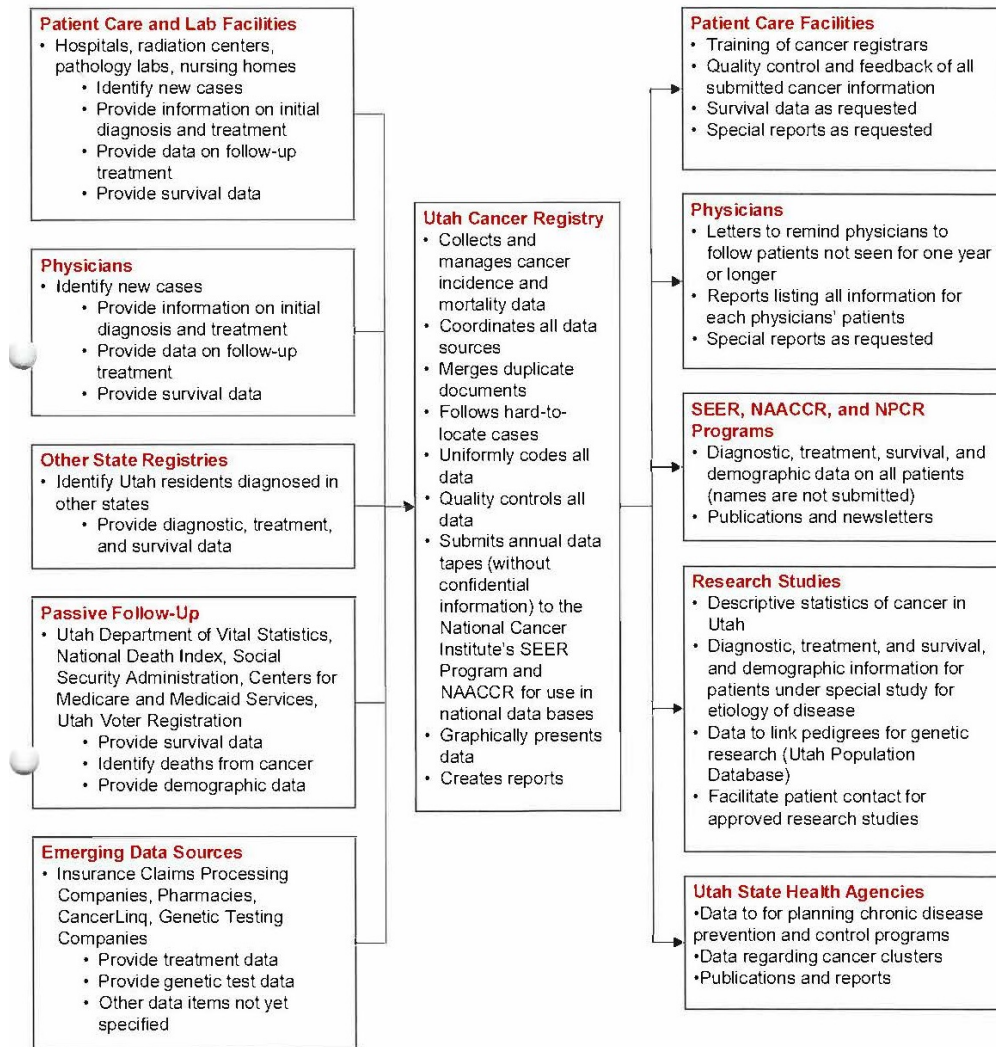
(19) UCR のデータ公表（実際）

- UDOH
- NCI SEER Program
 - SEER を通して、他のサーベイランスや調査での利用も
- CDC NPCR
- Utah Population Database
- 研究者からの要請
- 研究者以外からの要請

(20) UCR におけるデータ処理の流れ

次ページ

Utah Cancer Registry Data Flow



(21) Utah Cancer Registry People and Organization

UCR では、中央がん登録室として十分に機能するため、いくつかの鍵となる領域での専門家がいます。

医師

- 疫学及び研究方法
- Oncology
- 情報科学

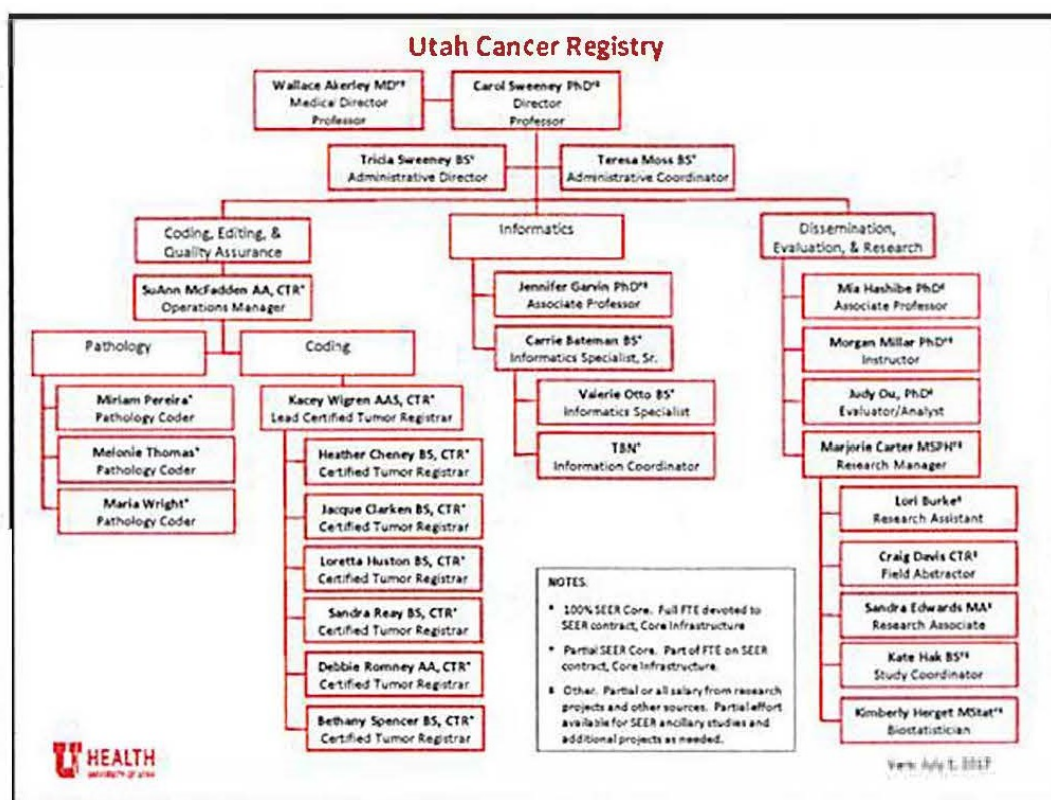
※ 以上の PhD、MD の有資格者は、UCR にパートタイムで来ており、残りの時間を大学での研究、教育、指導的役割に充てている。

他の有資格者

- Certified Tumor Registrar (CTR)
- Master Science in Public Health(MSPH)

- ❑ Master of Statistics(MStat)
- ❑ Bachelor of Information Science(BS)

(22) Utah Cancer Registry



- ※ CTR は総勢 8 人
- ※ Top は Director(Professor, PhD)で、Medical Director(Professor, MD)がサポート。
- ※ その Director の下に、Administrative Director(BS)と Administrative Coordinator(BS)が配置。
- ※ 実働部門として、① Coding, Editing & Quality Assurance、② Informatics、③ Dissemination, Evaluation & Research の 3 つの部門が設置。
- ※ Coding, Editing & Quality Assurance 部門には合計 11 人が配属され、Operations Manager (CTR) が Top で、その下に Coding と Pathology Coding の 2 部門があり、7 人の CTR が Coding を担当し (Lead CTR が 1 人)、Pathology の Coding を 3 人の Pathology Coder が担当。
- ※ Informatics 部門には合計 4 人が配属され、最上席は Associate Professor(PhD)で、次席にシニアの Informatics Specialist(BS)が就き、その Informatics Specialist の下に Informatics Specialist(BS)1 人と Information Coordinator が配属。
- ※ Dissemination, Evaluation & Research 部門には合計 9 人が配属され、最上席は Associate Professor(PhD)で、次席に Instructor(PhD)、Evaluator/Analyst(PhD)、Research Manager (MSPH)が就き、その Research Manager の下に Research Assistant、Field

Abstractor(CTR) 、 Research Associate (MA) 、 Study Coordinator(BS) 、 Biostatistician(MStat)の 5 人が配属。

2 Data Standards, Standard Setters & Supporting Organizations

(1) データの標準

- 異なったグループ間に一貫性を提供
- 米国で時を越えて発展
- 次の事項を決定する
 - 何を集めるのか
 - どのようにデータ処理を行うのか
 - どのようにデータを利用するのか
- データとデータ管理に関する質問に答えるために一番重要なこと

(2) つづき

- 共有された標準
 - コミュニケーションの透明度を確実にする
 - 蓄積されたデータや比較データの完全性を守る
 - がんのケア・がん対策の主要な側面に焦点を絞る

(3) つづき

- コンピュータ化された登録データ・システムがデータの共有を促進している
 - データ抽出のための NAACCR のレイアウト・フォーマット
 - ※ SEER も NCPR もこのフォーマットを利用している。

(4) 標準の設定

- 腫瘍の局在、形態及び進展度に関する標準コード
 - World Health Organization(WHO)
 - ・ ICD-O-3 部位・組織コード "Purple Book"
 - ・ コード選択のルール
 - American Cancer Society(ACS)
 - ・ 早期コーディングシステム; WHO の ICD-O に先行
 - American Joint Committee on Cancer(AJCC)
 - ・ Tumor-Node-Metastasis(TNM) prognostic system (※TNM system)
 - ・ International Union against Cancer(UICC)との協力

(5) つづき

- 施設あるいはコミュニティーベースの登録における標準
 - Commission on Cancer of the American College of Surgeons (CoC)
 - ・ がん対策における施設の役割の明確化
 - ※ Cancer Hospital の認証を行っている。
 - ・ National Cancer Data Base(NCDB)

- National Cancer Institute's Surveillance, Epidemiology & End Results Program(SEER)
 - ・ がんのモニタリングのための中央登録室の手続き
- National Program of Cancer Registries(NPCR)-Centers for Disease Control & Prevention(CDC)
 - ・ ポピュレーションベースの登録のためのガイドライン
- ※ 上記の標準の話とは別に、NCI の national cancer center(NCI-Designated Cancer Centers)が cancer center program のもとで指定されており、全米で 69 施設が指定されている。

(6) 支援組織

- The National Cancer Registrars Association(NCRA)
 - Cancer Registrars の支援 (※トレーニングなど)
- The North American Association of Central Registries(NAACCR)
 - 標準の開発
 - 標準の設定の支援
 ※ data exchange program や training webinar の開催など
- Utah Cancer Registrars Association
 - 中央登録室と院内がん登録を支援するための地方組織
 ※ お互いに顔見知りで、中には同じところで 10 年以上働いている知り合いもいて、お互いに何を期待しているのかを知っていて、助け合っている。

 ※ お互いの連絡には、e-mail、電話が使われている。

 ※ 10 月に年に 1 回のミーティングが開催される。

 ※ 大きな州の組織とは異なり、お金のない“Shoestring organization”である。

(7) 報告のための必要要件

- 施設の監督官庁に基づく
- 対象症例の重複
 - Utah Department of Health/Utah Cancer Registry
 - Utah CoC approved facilities
 - Hospital Cancer Committees
 - 'Reportable by Agreement' cases
 - 病院は UCR の要求に応えなければならない
 - UCR は SEER に要求されていない症例も受け入れている
 ※gastrointestinal stromal tumor(GIST) など

3 Data Collection, Training & Hospital Collaboration: What We Collect

(1) データの収集

- データの収集地域はユタ州の全住民を含む

- 院内での登録によって提出されれば、州外の症例はデータベースに含まれる
 - 漏れのない登録のために拡張されたケース・ファインディングが行われている
 - 早期の前立腺がん、皮膚のメラノーマ、同定困難な CML
 - ※そこで、urologist や dermatologist を対象に調査を行い、また、検査室にも調査を行っている。
- (2) SEER の対象症例
- 浸潤性及び in situ の新生物
 - 1996/1/1 から子宮頸がんの in situ は報告対象外に
 - 2004/1/1 以降に診断された脳・中枢神経系の良性腫瘍
 - 皮膚の基底細胞がん と扁平上皮癌は除外 (※容易に治療され、小さく、数が多すぎるため)
- (3) がんの症例の情報源
- ユタ州の 48 病院
 - “NAACCR Abstract”のフォーマット (※) で提出しているがん登録室のある病院
 - “フィールド・ホスピタル” (※州の北部にある小規模病院などで、registrar がおらず、registry もない)
 - 7つの独立した放射線センター
 - 医師のオフィス及びクリニック (※paper form での提出 (FAX)。Form は Web からダウンロード可能。ケース・ファインディングの意味合いで実施。実際の届出はそれほど多くはなく、約 10~1,000 件程度)
 - 病理検査室 (病院及び独立)
 - 死亡診断書
 - ナーシング・ホーム
 - 他の州とのデータ交換
- (4) 病理報告の収集
- 電子データでの病理報告
 - 88%が e-Path (※) を通して受理
 - (※NPCR-AERRO ePath Reporting Initiatives のこと。CDC が資金を提供している。ユタ州は検査室などとの間で NCI によるシステムを展開していたため、CDC による e-Path 導入が遅れた。)
 - ・ 48 病院中 36 病院
 - ・ 5つの独立した検査室
 - ・ AIM 社を採用している他の検査室の協力
 - Artificial Intelligence in Medicine (AIM 社)
 - 電子データ以外による病理報告
- (5) つづき

- 病院の病理検査室
 - 病理検査室をもつ 28 病院
 - ・ いくつかの検査室は複数の病院の path を読み取ることが可能
- 独立した病理検査室
 - 45 のユタ州内外の検査室
 - 21 施設は定期的に症例を提出
 - 他は散発的に提出
- 病理検査室の完全性
 - 検査室を見極めるために、調査員を派遣し、皮膚科医、泌尿器科医、腫瘍内科医に対して調査

(6) 病理検査のレビュー

- それぞれの病理レポートは報告対象かどうか決定するため UCR の職員によるレビューを受ける

(7) 病理検査の振り返り調査手順

- 病理報告が受理され、病院からの抽出されたデータとリンクしなかった場合には、「follow back」レターが path report に掲載された医師（※病院）に送られる。（※年末に実施。数百件ほど）
- 詳細な情報についての依頼
 - path report に記載されたがんについて入院したかどうかの確認
 - 診断情報
 - 治療

(8) 病院からの抽出情報の収集

- すべての病院の抽出情報は電子的に提供される（※NAACCR のフォーマットで）
 - 我々の保証された FTP サイトを通しての提供
 - SEER*DMS（※）抽出ツールからの転送
 - ※ SEER*DMS…SEER Data Management System
- 33 の病院は、病院の登録室から、または、抽出を行う会社と契約して報告
- UCR は、15 病院については抽出を行っている
- 診断から 6 か月以内に提出を依頼

(9) 記録の種類別での受理件数

記録の種類		年間平均の受理件数 (2012-2015 年)
HL7-E-Path（※）	←病理レポート	44,466
症例抽出(E-path 以外)	←病理レポート	4,035
NAACCR Abstract	←病院の記録	15,083
死亡診断書		17,293

※ HL7-E-Path…e-Path のこと

※ ユタ州の年間罹患数は約 10,000 件なので、E-Path の受理件数が年間平均だが 44,000 件だと 4 倍以上ということになる。これは、Unreportable が含まれることなどによる。最終的にはデータの Consolidation が重要なプロセス。

(10) 死亡診断書の整理手順

- 毎年、ユタ州の人口動態記録部門から死亡者ファイルの提供を受ける
- 登録データベースとマッチングを行う
 - 名前、ソーシャル・セキュリティー・ナンバー、出生日でマッチング
- マッチしたケースをアップデートする
 - 死亡日、死因
 - 出生地、人種、名前
- マッチしなかったケースは、“Follow back”により追加情報が届くまでは、“death certificate only(DCO)”の可能性ありの症例と指定

(11) DCO 症例に対する Follow-Back の手順

- 病院での死亡
 - 抽出情報が提出されたかどうか確認するため病院にコンタクトする
- ナーシング・ホームでの死亡
 - 情報収集のためコンタクトをとる
- 他の場所での死亡
 - 死亡診断書にサインした医師に対して手紙を送る
 - 死亡診断書に記載された情報提供者に手紙を送る
 - ・ Death clearance guideline (※) に従い、限定された情報を利用
 - ※ Death clearance guideline…NAACCR による Death Clearance Manual
 - ・ 時間経過と診断した医師の情報提供を要求

(12) 死亡診断書に対する Follow-Back の手順

- 報告対象の確認
 - 診断時にユタ州の住民
 - 報告対象の診断名
- 指定された報告の情報源
 - 病院、医師のオフィス、DCO
- NAACCR の Death Clearance Manual により DCO ルールを正しく適用
- SEER は DCO を 0~1.5% で許容
- 2014 年の DCO は 0.37% (44 例)

(13) 2014 年の Death Clearance Report

- マッチしなかった死亡診断書 816 (※①~③の合計)
- UCR データベースに追加された症例 285 (※①)
 - 病院 8

- 放射線センター/オンコロジーセンター 4
 - 医師のオフィス 261
 - ナーシング・ホーム/ホスピス 5
 - 剖検のみ 7
 - DCO 症例 44 (※②)
- (14) つづき
- フォローが行われたが報告対象外 487 (※③)
 - すでに登録済みの症例 286
 - 対象地域外 117
 - 死亡原因が転移の部位 30
 - ・ 新しい原発部位ではない
 - 死因が報告対象外 54
 - ・ 報告対象外の皮膚がん

(15) UCR の症例 (2014 年・報告の情報源別)

情報源	症例数
1 病院の入院患者	10,293
2 放射線治療センター/オンコロジーセンター	15
3 病理報告のみ	454
4 医師のオフィス	1,252
5 ナーシング・ホーム	5
6 剖検のみ	8
7 DCO	44
8 他の病院外来ユニット/外科センター	4
合計	12,075

(16) トレーニング/質の管理 の手順

- 経験豊かで認証 (Certified Tumor registrar, CTR) を受けたエディターと質向上の専門家
- データを完全性、質、コーディング・ルールの正確な適用について評価
- 標準に合致しない抽出情報は、訂正のため元の施設に返却

(17) つづき

- ユタ州の病院の登録者とは仕事上良好な関係にある (※電話、メールなどで連絡し合っている)
 - 技術的な支援を提供
 - ユタ州のがん登録者協会のミーティング、NAACCR や NCRA の webiner (ウェブカンファランス)、SEER*Educate を通して、トレーニングの機会を提供
- ※ NAACCR Webinars-Series…<https://education.naacccr.org/Webinars-Series>

- ※ NCRA Webinar···<https://www.ncra.org/Certifications/content.cfm?ItemNumber=8631&navItemNumber=11709>
- ※ SEER*Educate ···<https://educate.fredhutch.org/LandingPage.aspx>

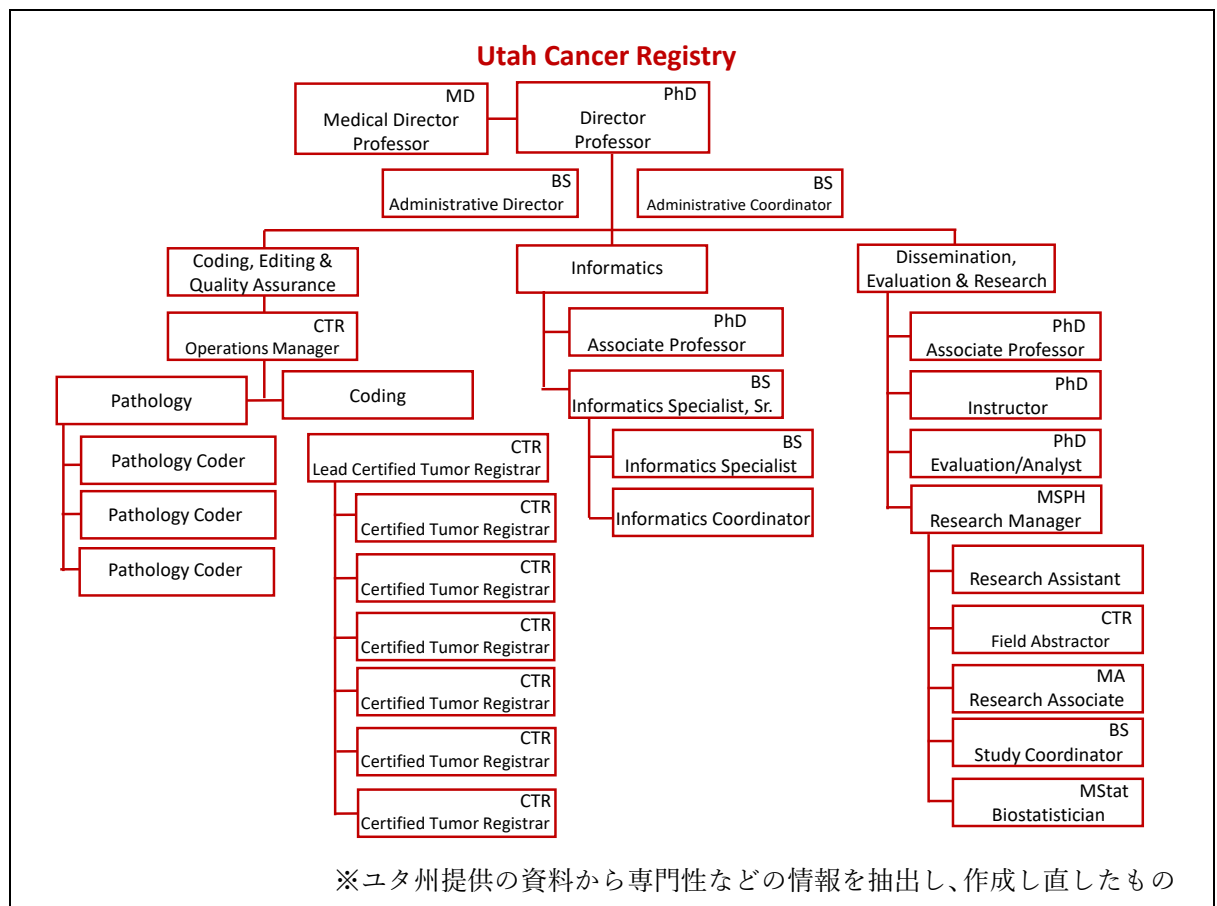
米国における情報収集と提供の仕組み（ユタ州がん登録を例として）

1 ユタ州のがん登録

(1) 概要

- ・ ユタ州全域を対象とした population-based のがん登録
 - ・ 1966 年からユタ大学で運営
 - ・ 1973 年から SEER プログラムに参加（NCI との契約に基づく→Registrar の雇用）
 - ・ 2017 年 6 月から NPCR に参加（CDC との協力協定に基づく→Registrar の雇用）
- ポイント①データ集約・データセット作成を行う機関では、外部資金によるスタッフの継続雇用への方策がとれないか？

(2) スタッフ



- ・ CTR は総勢 9 人。Top は Director(Professor, PhD)で、Medical Director(Professor, MD)がサポート。その Director の下に、Administrative Director(BS)と Administrative Coordinator(BS)が配置。
- ・ 実働部門として、①Coding, Editing & Quality Assurance、②Informatics、

③Dissemination, Evaluation & Research の3つの部門が設置。

- ・ Coding, Editing & Quality Assurance 部門には合計 11 人が配属され、Operations Manager (CTR) が Top で、その下に Coding と Pathology Coding の 2 部門があり、7 人の CTR が Coding を担当し (Lead CTR が 1 人)、Pathology の Coding を 3 人の Pathology Coder が担当。
- ・ Informatics 部門には合計 4 人が配属され、最上席は Associate Professor(PhD)で、次席にシニアの Informatics Specialist(BS)が就き、その Informatics Specialist の下に Informatics Specialist(BS)と Information Coordinator が配属。
- ・ Dissemination, Evaluation & Research 部門には合計 9 人が配属され、最上席は Associate Professor(PhD)で、次席に Instructor(PhD)、Evaluator/Analyst(PhD)、Research Manager (MSPH)が就き、その Research Manager の下に Research Assistant、Field Abstractor(CTR)、Research Associate (MA)、Study Coordinator(BS)、Biostatistician(MStat)の 5 人が配属。

→ポイント②データ集約・データセット作成を行う機関では、収集部門と情報提供部門に分けて、それぞれに専門職種のマネージャーを配置できないか？

- ・ Advisory Research Committee
 - 年 2 回の会合に、UDOH (Utah Department of Health)、病院、研究者からの代表が参加

(3) データ収集と提供等

- ・ 全体の流れとしては、10~12 ページのとおり
- ・ データの収集
 - 対象はユタ州の全住民
 - 院内での登録によって提出されれば、州外の症例も登録
 - 漏れのない登録のために拡張されたケース・ファインディング (早期の前立腺がん、皮膚のメラノーマ、同定困難な CML を対象に泌尿器科医・皮膚科医・検査室にも調査)

情報源

- ユタ州の 48 病院

- ❑ NAACCR Abstract のフォーマットで提出可能ながん登録室のある病院と registrar がいない州北部にある小規模病院 (field hospital)
- ❑ SEER*DMS (SEER Data Management System) という抽出ツール利用し、UCR の保証された FTP サイトを通して提供
- ポイント③DPC 対象病院などに対して、病名などから対象症例を抽出するツールを開発したり、院内がん登録全国集計への参加病院を対象に抽出・登録用のツールを開発する方策がとれないか？
- ❑ 33 病院は病院の登録室から直接、または、抽出を行う会社と契約して報告され、15 病院については、UCR が抽出を行っている。
- ❑ 診断から 6 か月以内に提出を依頼

- 7つの独立した放射線センター
- 医師のオフィス及びクリニック（泌尿器科医・皮膚科医）
 - 紙様式で FAX での提出、case finding の意味合いで実施
- 病理検査室（28 病院及び 45 の州内外の独立した検査室）

□ 88%が e-Path を通して受理

（注）CDC が資金を提供している NPCR-AERRO ePath Reporting Initiatives によるシステムによる e-Path ではない！ユタ州は検査室などとの間で NCI（SEER）によるシステムを展開しており、ライセンスの問題があり、CDC による e-Path 導入が遅れ、NPCR への参加が遅れた。

□ AIM 社による e-path を採用している検査室から（13～15 ページ参考）

➔ ポイント④病理報告に関する様式の標準化を進めることが出来ないか？

➔ ポイント⑤標準化された病理報告を電子データ化し、収集できる仕組みが
出来ないか？

□ 病理検査の完全性を確認するため、調査員を派遣し、調査を行っている。

□ 病理報告は、報告対象かどうか UCR の職員によるレビューを受ける。

□ 病理報告受理後、振り返り調査を行う。

➢ 病理報告が受理され、病院からの抽出されたデータとリンクしなかった場合には、「follow back」レターが path report に掲載された医師（※病院）に送られる（年末に実施、数百件ほど）。

➢ 詳細な情報提供について依頼（path report に記載されたがんについて入院したかどうか、診断情報、治療情報）

- 死亡診断書

□ 毎年、ユタ州の人口動態記録部門から死亡者ファイルの提供を受ける。

□ 登録データベースとマッチングを行い（名前、ソーシャル・セキュリティー・ナンバー、出生日）、データをアップデート（死亡日、死因、出生地、人種、名前）。

□ マッチしなかったケースには Follow-Back 調査を実施

➢ 病院での死亡➔抽出情報が提出されたかどうか病院に確認

➢ ナーシング・ホームでの死亡➔情報収集のためコンタクトをとる

➢ 他の場所での死亡➔死亡診断書にサインした医師、情報提供者に手紙を送る。

□ NAACCR による Death Clearance Manual に従い、限定された情報を収集・利用

（つづく）

□ 2014 年の Death Clearance Report

区分	件数
データベースに追加	285
うち、病院	8
放射線センター/オンコロジーセンター	4
医師のオフィス	261
ナーシング・ホーム/ホスピス	5
剖検のみ	7
報告対象外	487
うち、すでに登録済み	286
対象地域外	117
死亡原因が転移の部位（新しい部位ではない）	30
死因が報告対象外（対象外の皮膚癌※）	54
DCO	44
合計（マッチしなかったケース）	816

※基底細胞癌と扁平上皮癌（※容易に治療され、小さく、数が多すぎる）

- ナーシング・ホーム
- 他の州とのデータ交換

・ 届出等の受理件数（2012-2015 年）

記録の種類	年間平均の受理件数
HL7-E-Path（※病理報告）	44,466
症例抽出(E-path 以外（※病理報告）	4,035
NAACCR Abstract（※病院からの記録）	15,083
死亡診断書	17,293

・ UCR の症例（2014 年・報告の情報源別）

情報源	症例数
1 病院の入院患者	10,293
2 放射線治療センター/オンコロジーセンター	15
3 病理報告のみ	454
4 医師のオフィス	1,252
5 ナーシング・ホーム	5
6 剖検のみ	8
7 DCO	44
8 他の病院外来ユニット/外科センター	4
合計	12,075

※2014 年の DCO は 0.37%（SEER は DCO を 0~1.5%で許容）

- ・ データの提供
 - SEER と NPCR には、個人同定できない形式でデータを提出
 - SEER と NPCR は、がん登録からのデータを集めて大きなデータセットにし、がん罹患のトレンド評価やがん研究に利用可能としている。
 - ➔ ポイント⑥SEER や NPCR を参考に、がん登録室から集めた非匿名化情報を研究利用可能とする制度設計は可能か？（SEER 訪問結果は？）
 - SEER(15 州が参加)はがん研究へのデータ提供(最終的な目標は survival)、NPCR (46 州が参加) はがん対策へのデータ提供というスタンスの違いがある。
- ・ データの公表
 - Utah Department of Health(UDOH)
 - 登録データは、UDOH が所有。
 - UCR は 1 年分のデータを UDOH に送り、概要報告を提供し、公衆衛生の目的のために UDOH とのデータリンケージを行う。
 - UDOH とユタ大学との合意覚書により、他の権限付与されたデータの公表について明確化されている。
 - UCR の指導者は月 1 回 UDOH と会合を行っている。
 - NCI SEER Program
 - SEER を通して、他のサーベイランスや調査での利用もある
 - ➔ ポイント⑦SEER を通した二次利用の方法を参考に、日本でも仕組みづくりができないか？
 - CDC NPCR
 - Utah Population Database
 - 研究者からの要請
 - 研究者以外からの要請

3 米国における情報収集と提供の仕組み

(1) 仕組み

① 法的な枠組みの有無と内容（ユタ州）

- ・ 州法：Utah Code Title 26-5-3 (Title 26 Utah Health Code, Chapter 5 Chronic Disease Control, Section3, System for detecting and monitoring diseases established by department)

❑ The department (of health) shall develop and maintain a system for detecting and monitoring chronic diseases within the state and shall investigate and determine the epidemiology of those conditions which contributed to preventable and premature sickness, or both, and to death and disability.

- 行政規則：Utah Administrative Code Rule R384-100 (Cancer Reporting Rule)

- ❑ All hospitals, radiation therapy centers, pathology laboratories licensed to provide services in the state, nursing homes, and other facilities and health care providers involved in the diagnosis or treatment of cancer patients shall report or provide information related to a cancer or reportable benign tumor to the Registry.
- ❑ Cancer records are managed by the Utah Cancer Registry (Registry) on behalf of the Utah Department of Health.
- ❑ The Utah Department of Health retains ownership and all rights to the records.
- ❑ All reports required by this rule are confidential under the provisions of Title 26, Chapter 3 and are not open to inspection except as allowed by Title 26, Chapter 3. The Registry shall maintain all reports according to the provisions of Title 26, Chapter 3.”

- ユタ州の経験からは

- ❑ このルールは、データの形式や変数のような具体にまでは言及していない。
- ❑ 言い回しが具体的ではないため、SEER、NPCR、NAACCR などが登録における変数やデータ形式に変更を要求してきても、ルールの変更を必要としない。

- 州内の病院に対する Reporting Rule の効果と実際

- ❑ UCR への報告のために個々の症例について詳細な情報を収集することを目的として、病院は cancer registrar を雇用する、あるいは、このようなサービスを提供する会社と契約している。病院はこの「データの抽出」の費用を支払わなければならない。
➡ポイント⑧全国がん登録にない項目を抽出する方法として、このようなデータ収集を行う機関・団体ないし会社を利用する方法がとれないか？
➡ポイント⑨その場合の機関・団体ないし会社へ支払う費用は、研究を維持していく仕組みの中で賄うことができなければ、研究は維持できなくなる。つまり予算・財源の目途をつけることが必要。
- ❑ ユタ州の病院やヘルスケア提供施設は、全般的に報告に対してとても協力的である。UCR は日常的に病院やヘルスケア提供施設と接触し、タイムリーに報告するよう奨励している。UCR とユタ州はこれまで登録規則の法的な強制力を行使したことはない。

- ② 全米での仕組みの活用（データの標準化、施設認定、ガイドライン）
- ・ ICD-O-3(American Cancer Society(ACS) & WHO)
 - ・ TNM system(American Joint Committee on Cancer(AJCC) & UICC)
 - ・ データ抽出のために NAACCR abstract のフォーマットを利用(SEER も NCPR もこのフォーマットを利用)
 - ・ Commission on Cancer of the American College of Surgeons (CoC)による National Cancer Data Base(NCDB)
 - ・ SEER によるがんのモニタリングのための中央登録室の手続き
 - ・ CDC による NPCR のためのポピュレーション・ベースの登録のためのガイドライン

→ポイント⑩全国がん登録にない項目を抽出する場合であっても、データの標準化などの基本的な部分を担保する役割を誰が担うのか整理すべき

- ③ 資金的援助を伴うプログラムとしての枠組み（SEER、ユタ州）
- ・ SEER や NPCR に参加したがん登録は、登録データの標準化と品質保証のための特別プロジェクトに参加申請をする機会（以下に例示）、メールでの情報提供やトレーニングの機会（SEER*Educate）などが得られる。

- SEER*Educate …<https://educate.fredhutch.org/LandingPage.aspx>
- SEER Virtual Tissue Repository
- 15-month Resubmission to Improve Completeness
- Pilot to Collect Screening Information for Breast Cervical Cancer Cases
- UCR Data Linkage with Utah All-Payer Claims
 - 日本と保険制度が異なり、保険が変わることもあるため重要
 - データリンケージは UCR で行っていて、Social Security Number があるため、リンケージは高い確率で一致。Social Security Number が提供してもらえないときは、リンケージには誕生日と名前を利用するが、ラストネームが変わったり、ニックネームや略称などがあるため、困難を伴うことがある。

- ・ NAACCR による標準の開発と支援

- Central registries(population-based registries)が参加し、重要な役割をはたしている。
- data exchange program や training webinar の開催などを行っている。
- Webinars-Series…<https://education.naacr.org/Webinars-Series>

- ・ The National Cancer Registrars Association(NCRA)による Cancer Registrar の支援

- ❑ Cancer Registrar が各個人で参加。
- ❑ トレーニングの機会（有料）を提供し、Certified Tumor Registrar (CTR) の認証を行っている。
- ❑ トレーニングはいくつかのコースがあり、期間が異なる。通常、2年間の実務期間の後に認証の試験を受ける。）
- ❑ Webinar…<https://www.ncra.org/Certifications/content.cfm?ItemNumber=8631&navItemNumber=11709>
- ❑ 州によって違いはあるかもしれないが、CTR の認証を受けることは雇用に重要となっている。ユタ州では、実際に入職する前にもトレーニングを受けている。

・ Utah Cancer Registrars Association

- ❑ 中央登録室と院内がん登録を支援するための地方組織
- ❑ お互いに顔見知りで、中には同じところで10年以上働いている知り合いもいて、お互いに何を期待しているのかを知っていて、助け合っている。
- ❑ お互いの連絡には、e-mail、電話が使われている。
- ❑ 10月に年に1回のミーティングが開催される。
- ❑ 大きな州の組織とは異なり、お金のない“Shoestring organization”である。

→ポイント⑩全国がん登録にない項目を収集する場合、病院の協力を得る必要がある場合には、情報提供やトレーニングの機会の提供が必要

(2) 必要経費 →今回はデータの入手までには至っていない

- ① 経費の総額・内訳
- ② 資金的援助 (SEER)
- ③ 資金的援助 (ユタ州)
- ④ 資金的援助 (他機関等)

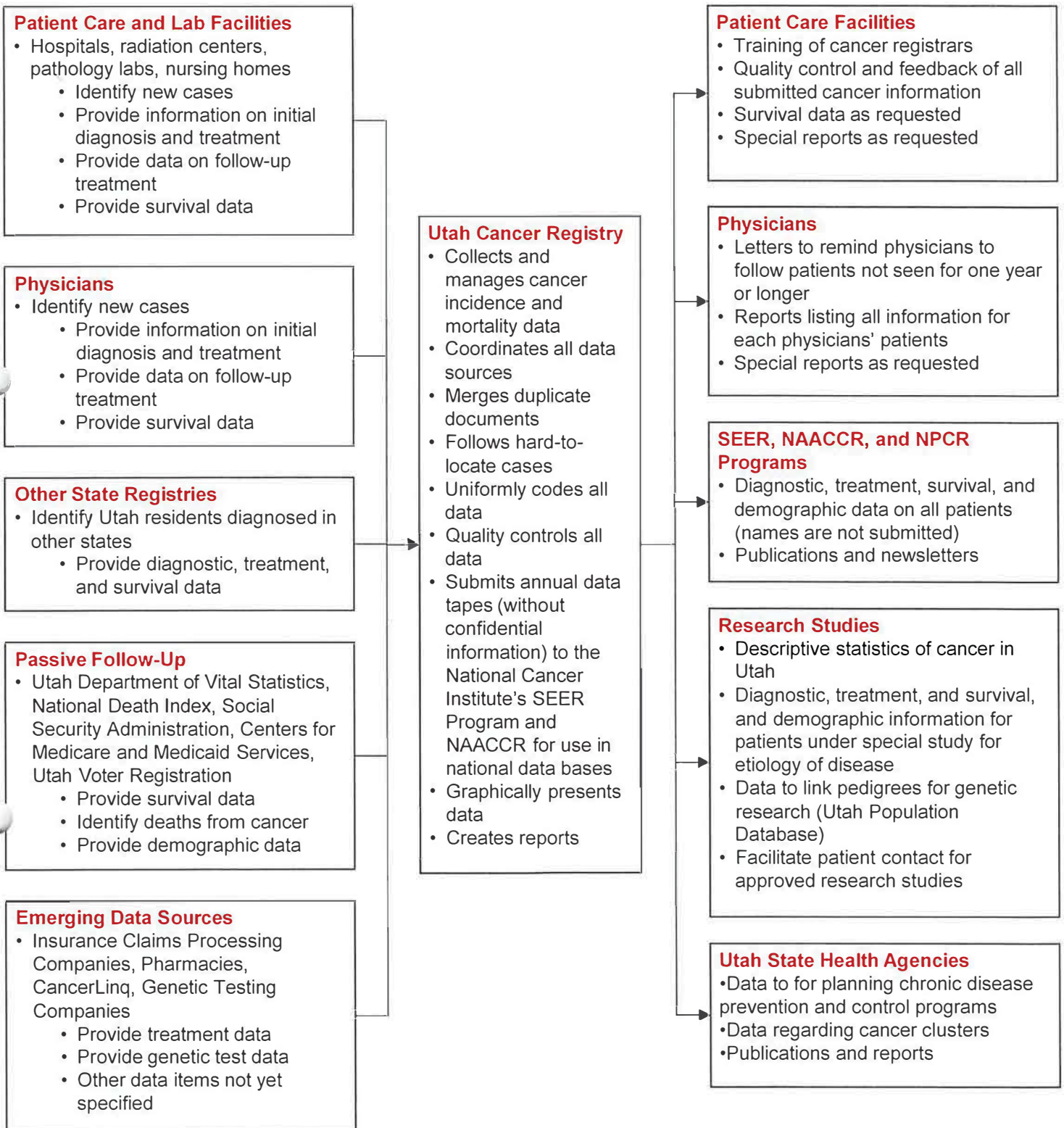
(3) 教育支援 →上記 (1) ③で掲載

- ① 教育プログラム
- ② 技術的支援
- ③ その他

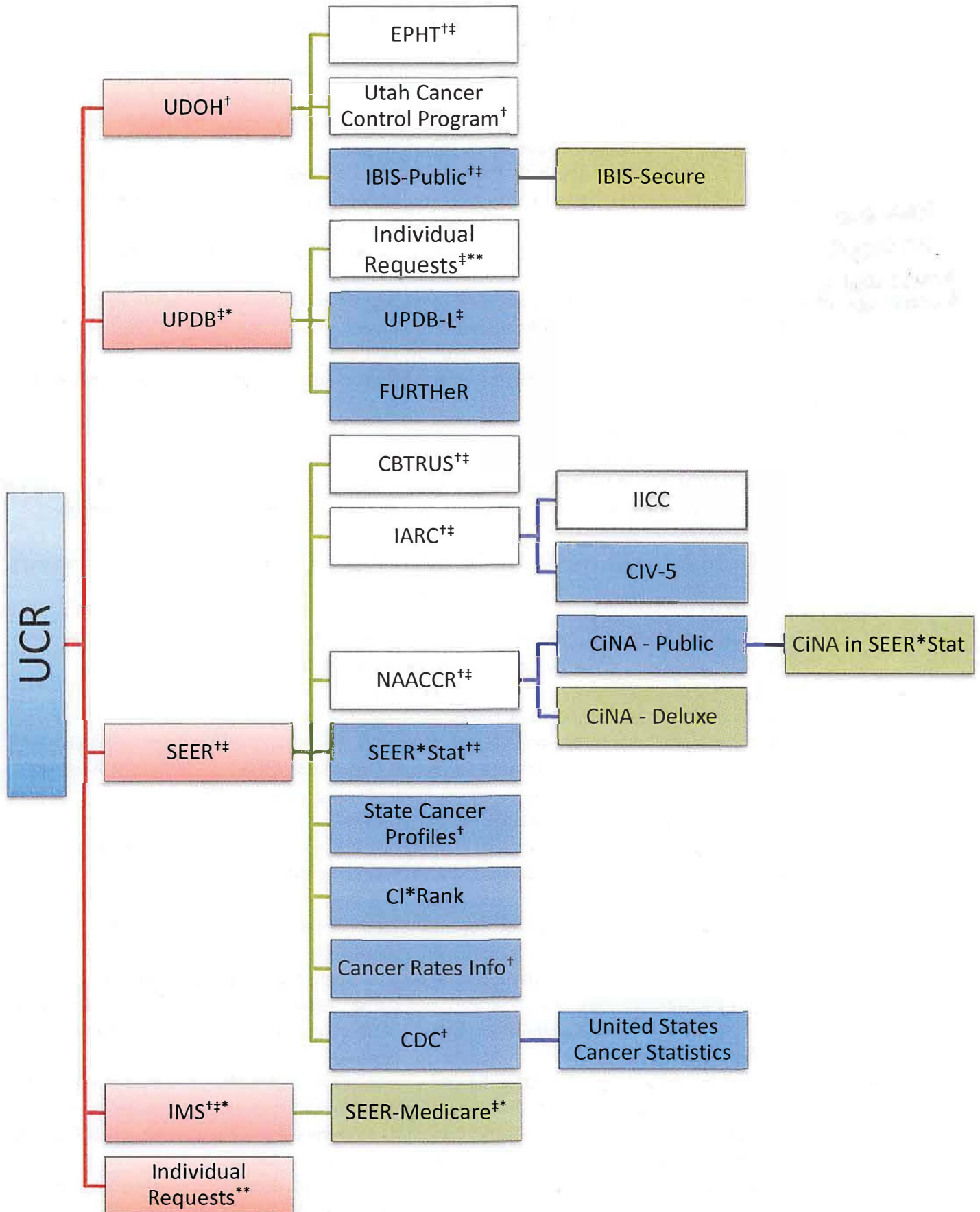
4 ポイントのまとめ（提言に向けて）

- ① データ集約・データセット作成を行う機関では、外部資金によるスタッフの継続雇用への方策がとれないか？
- ② データ集約・データセット作成を行う機関では、収集部門と情報提供部門に分けて、それぞれに専門職種のマネージャーを配置できないか？
- ③ DPC 対象病院などに対して、病名などから対象症例を抽出するツールを開発したり、院内がん登録全国集計への参加病院を対象に抽出用のツールを開発する方策がとれないか？
- ④ 病理報告に関する様式の標準化を進めることが出来ないか？
- ⑤ 標準化された病理報告を電子データ化し、収集できる仕組みができないか？
- ⑥ SEER や NPCR を参考に、がん登録室から集めた非匿名化情報を研究利用可能とする制度設計は可能か？（SEER 訪問結果は？）
- ⑦ SEER を通した二次利用の方法を参考に日本でも仕組みづくりができないか？
- ⑧ 全国がん登録にない項目を抽出する方法として、このようなデータ収集を行う機関・団体ないし会社を利用する方法がとれないか？
- ⑨ その場合の機関・団体ないし会社へ支払う費用は、研究を維持していく仕組みの中で賄うことができなければ、研究は維持できなくなる。つまり予算・財源の目途をつけることが必要。
- ⑩ 全国がん登録にない項目を抽出する場合であっても、データの標準化などの基本的な部分を担保する役割を誰が担うのか整理すべき
- ⑪ 全国がん登録にない項目を収集する場合、病院の協力を得る必要がある場合には、情報提供やトレーニングの機会の提供が必要

Utah Cancer Registry Data Flow



Utah Cancer Registry Data Dissemination Chart



† Surveillance, ‡ Research, * Identifiable data, **May contain identifiable data

Red=Initial UCR Data Release, Blue=Publically Available Queries, Green=Query systems that require approval

UDOH – Utah Department of Health. Governed by a memorandum of agreement, UDOH receives a full dataset, which is used for surveillance and research. (Identifiability: full dates and addresses, but no names.)

- EPHT – Environmental Public Health Tracking. This UDOH program uses UCR data for cancer cluster investigations and to provide reports to various CDC programs, including the national EPHT program. (<http://epht.health.utah.gov/epht-view/>)
- Utah Cancer Control Program – information on cancer screenings performed in Utah are linked to UCR data for public health surveillance purposes. (<http://cancerutah.org/>)
- IBIS-PH – Indicator-Based Information System for Public Health (<http://ibis.health.utah.gov/>). This UDOH Bureau of Health Promotion program provides a query system for cancer, and other health topics, available to the public.
 - IBIS Public: de-identified, queryable by cancer site, diagnosis year, demographic variables
 - IBIS Secure: Includes zip code and census tract. Users must go through a review process similar to an IRB to gain access

UPDB – Utah Population Database. Governed by a memorandum of understanding, UPDB receives a full dataset

<https://healthcare.utah.edu/huntsmancancerinstitute/research/updb/>

(Identifiability: They receive full identifying information (Name, DOB, SSN, etc.))

- Individual requests – Data released to individual researchers after approval by IRB and RGE. May or may not contain identifiable data.
- UPDB-L – this data is public, de-identified, and queryable (<https://hci-updblapp.hci.utah.edu/updbl/>)
- FURTHeR – U of U Query Tool, same data as UPDB-L, limited to U of U patients (<http://www.further.utah.edu/>)

SEER – The Surveillance Epidemiology and End Results program. Governed by a contract, primary source of UCR funding, part of National Cancer Institute (Identifiability: They receive full dates for DOB/DOD/DODx, but no other identifiers.)

<https://seer.cancer.gov/>

- CBTRUS – Central Brain Tumor Registry of the US – They receive benign brain tumor cases (<http://www.cbtrus.org/>)
- IARC – International Agency for Research on Cancer - this data is used for surveillance and research (<https://www.iarc.fr/>)
 - IICC – International Incidence of Childhood Cancer, Publication (<http://iicc.iarc.fr/>)
 - CIV-5 – Cancer Incidence in Five Continents, Public and queryable (<http://ci5.iarc.fr/>)
- NAACCR – North American Association of Central Cancer Registries
 - Cancer in North America Public (CiNA) data sets
 - CiNA Public – limited variables, requires DUA (<https://www.naaccr.org/cina-public-use-data-set/>)
 - CiNA in SEER*Stat – needs NAACCR approval to access (<https://www.naaccr.org/cina-in-seerstat/>)
 - CiNA Deluxe –used for research, requires full application/review for access (<https://www.naaccr.org/cina-deluxe-for-researchers/>)
- SEER*Stat –data set is public and queryable after signing a confidentiality form (<https://seer.cancer.gov/seerstat/>)
- State Cancer Profiles –data set is public and queryable, very limited query options (<https://seer.cancer.gov/statistics/scp.html>)
- CI*Rank – Cancer incidence and mortality, queryable by county, cancer site, diagnosis year, demographics (<https://surveillance.cancer.gov/cirank/>)
- Cancer Rates Info –data set is public and queryable. Data through 2009 only (<http://www.cancer-rates.info/>)
- Centers for Disease Control and Prevention – cancer data collected by SEER registries and CDC’s National Program of Cancer Registries are combined to produce the official federal statistics on cancer incidence. Deidentified data sets can be downloaded for analysis. (<https://nccd.cdc.gov/uscs/>)

IMS – Information Management Systems. IMS contracts with NCI to develop and manage the Database Management System (SEER*DMS) for SEER registry data. IMS hosts Utah Cancer Registry data. UCR staff add, edit, and update cancer case information in SEER*DMS, working on the IMS-hosted data set via an encrypted virtual private network (VPN) in real time (Identifiability: UCR sends identifiers to IMS) <https://www.imsweb.com/products/>

- SEER Medicare – this data is used for research and is restricted. IMS links and sends Medicare data to SEER, who can link using a SEER Medicare ID (<https://healthcaredelivery.cancer.gov/seermedicare/>)

Individual Requests - Data released to individual researchers as requested and approved by UCR, IRB, RGE, and/or ARC committees (committee review depends on nature of request). May or may not contain identifiable data.

Cancer Surveillance

E-Path Monitoring Console

Cancer Data Now

Reviewing electronic cancer reporting

The E-Path Monitoring Console provides submitting sites with several important functions to assist in the management of an automated E-Path system. These include:

- Process Monitoring
- Pathology Report Review
- HIPAA Disclosure Reporting

The following descriptions of these functions assume some familiarity with the E-Path technology provided by AIM. The E-Path Monitoring Console incorporates the functionality of AIM's ISIS Registrar, Coding Edition product. Information about E-Path can be made available to you at your request.

Process monitoring

E-Path is an automated process, designed to operate in an unattended manner. The process monitoring function provides, on demand, process reports on the operation of the system along with summary information about reports that have been processed or failed to process.

Summaries may be obtained for daily, monthly and annual periods.

Summaries tabulate the following:

- Total number of reports processed
- Total number of reports selected (as being reportable to the registry)
- Total number of reports that could not be processed
- Breakdown of the reports processed by report type (e.g. Pathology, Gynecological Cytology, Bone Marrow, Autopsy, etc.)

The count of pathology reports on summaries is not unique. Thus, if a report is sent to the system twice, for example, because of an amendment, it will be counted each time.

All pathology reports that could not be processed are placed in an error log. The E-Path Monitoring Console does not provide any function to retrieve these from the error log (they may be viewed using any text editor).

Processing summaries may be viewed on screen, printed or saved to text files.

Pathology report review and forwarding

The functionality of ISIS Registrar, Coding Edition, has been incorporated into the Console to provide it with the means to search for, view, forward and review the coding of pathology reports that have been selected by the E-Path filter.

Selected pathology reports are loaded into the Console's database. Individual reports may then be viewed by selecting them from a list of available reports.

Alternatively a search function is provided to find reports by various criteria, including diagnosis. Pathology reports are displayed in their entirety and in a standard format. They may not be edited. Pathology reports may be coded with the system's AutoCode function. The assigned topography and morphology codes are displayed and may be edited. The user may also select the codes, in the event there are several, under which the report is to be classified.

The coding system normally used by AutoCode is ICD-O-3. Other coding systems may be used, but this will require the system to be reconfigured by AIM.

Although E-Path is generally configured to automatically forward selected pathology reports to the designated central cancer registry, they may also be forwarded to other destinations determined by the users. In many cases, other destinations include the local (e.g. hospital or laboratory) cancer registry system and researchers within and external to the institution.

Pathology Reports may be exported in the following formats:

- NAACCR Abstract v 9.1 or 10.1**
- NAACCR HL7
- NAACCR Flat File

Exported pathology reports may be:

- Forwarded in patient identified or de-identified format
- Directed to a file
- Emailed

The selected pathology reports are stored in the Console's database and are available until the data are purged from the system. The design of the system assumes that pathology reports from several years will be on file at any one time.

HIPAA Disclosure Reports

The HIPAA Disclosure Report lists those pathology reports that have been automatically selected and forwarded from the institution to the central registry by the E-Path system. It will not include any pathology reports that are sent using the ISIS Registrar function to forward reports.

HIPAA Disclosure Reports may be viewed and printed for daily, monthly or annual periods. The user may select any or all of the following data elements to include in the report, provided they are available in the source records:

- MRN
- Accession Number
- Patient Name
- Patient Sex
- Patient Date of Birth
- Patient Account Number

E-Path HIPAA Patient Information Disclosure Report

Pleasant Valley Hospital

Report for the Period 2004-May-01 to 2004-May-31

MRN	Accession No.	Patient Name	Sex	Date of Birth	Date Disclosed	Disclosed To
MRN123456788	GP-10336	Jones, Kathryn	F	1970-12-21	2004-05-01 13:37:01	State Cancer Registry
MRN123456789	GP-10334	Kirk, Cameron	M	1963-08-10	2004-05-01 16:58:24	State Cancer Registry
MRN012234567	GP-11155	Armstrong, Michael	M	1955-02-15	2004-05-02 16:58:30	State Cancer Registry
MRN990123456	GP-90123	McDonald, Alley	F	1975-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN123456789	GP-10334	Frasier, Derek	M	1969-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN012234567	GP-11155	Baird, John	M	1970-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN990123456	GP-90123	Smart, Susan	F	1972-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN123456789	GP-10334	Croft, Teresa	F	1966-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN012234567	GP-11155	Appleton, William	M	1959-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN990123456	GP-90123	Whitley, Anthony	M	1958-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN123456788	GP-10336	Sean, Cynthia	F	1976-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN123456789	GP-10334	Brian, David	M	1977-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN012234567	GP-11155	Chow, Melanie	F	1980-05-08	2004-05-02 16:58:30	State Cancer Registry
MRN990123456	GP-90123	DeCruz, Grace	F	1976-05-08	2004-05-02 16:58:30	State Cancer Registry

E-Path Processing Summary

Pleasant Valley Hospital
Report for the Period of 2004-March-01 to 2004-March-31

Process Summary

Category	Total
Number of Reports Received from LIS:	702
Number of Reports Processed:	702
Number of Warnings:	0
Number of Reports Transmitted:	121
Number of Reports not Transmitted	581
Number of Reports not Processed:	0

Report Type Summary

Report Type	Number of Reports
Pathology	311
Cytology	91
Gynecological Cytology (Pap)	270
Bone Marrow	18
Autopsy	4
Other	8
Unknown	0
Total processed	702

- Social Security Number
- Date Disclosed

The recipient and reason for disclosure may also be specified and included in the report

Hardware Requirements and Installation

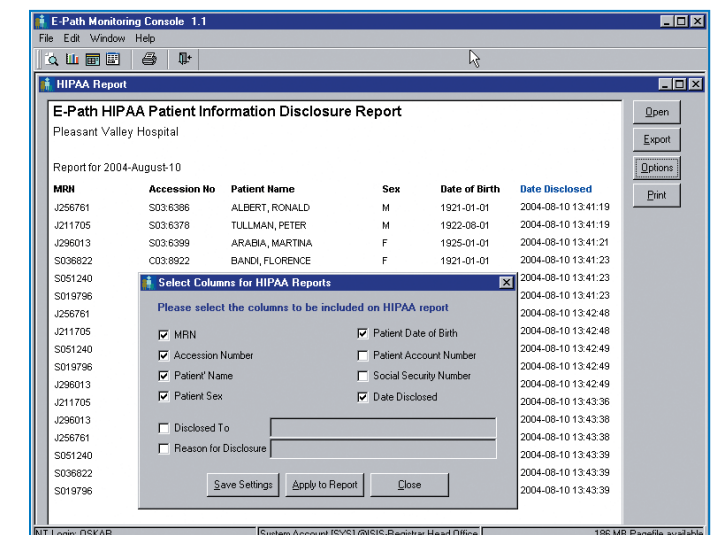
It is recommended that the E-Path Monitoring Console be installed on a different computer from the one on which E-Path itself is deployed. The two computers need to be connected to the same network.

The system is intended to run on a standard desktop PC running any recent version of Microsoft Windows.

Recommended specifications include:

- Intel Pentium II Processor
- 300 MHz or greater
- 64 MB RAM
- 2 GB available disc space

Note: The abstract record will contain only the components available in the pathology report.





2 Berkeley St., Suite 403
Toronto, Ont., Canada M5A 2W3
Tel: 416.594.9393 Fax: 416.594.2420
Toll Free 1.866.645.2224
E-mail: letstalk@aim.ca
Website: www.aim.ca

Our engineers and technicians understand the medical environment, its paradigms, processes and needs. AIM provides comprehensive information technology services to the healthcare industry to meet the special requirements of clinical and research systems.

Copyright © 2014, AIM Inc. All rights reserved. AIM, the AIM logo, are trademarks or registered trademarks of Artificial Intelligence In Medicine Inc.

The information contained herein is subject to change at any time without notice and does not represent a commitment on the part of AIM. AIM makes no warranties, express, implied or otherwise, with respect to this document. In no event shall AIM be responsible for any indirect, special, incidental or consequential damages arising out of the purchase or use of this document and/or the information/product described herein.

厚生労働科学研究費補助金（がん対策推進総合研究事業）
（分担）研究報告書

全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究
－仕組み検討（モデル地域候補）－

研究分担者 大木いずみ 栃木県立がんセンターがん予防情報相談部 部長

研究要旨

がん登録等の推進に関する法律（以下「がん登録推進法」）によってようやくわが国のがん罹患の推移を明らかにする仕組みができた。今後はいかにこれらのデータを活用し、公衆衛生活動や医療に生かしていくかが課題となる。全国がん登録の情報の利用と提供について様々な角度から検討されているが、米国の **Surveillance, Epidemiology, and End Results Program**（SEER 事業）のようなデータがわが国においても整備されれば、さらなる活用が進むことが期待される。がん登録推進法による新たながん登録が単なるデータ収集に終わらないよう、発展を遂げるために、研究用データを作成し **Study question** の解決につながる精度を保たなければならない。

今後の対策に活用するために、今年度は SEER 事業にデータを提供しているユタ州がん登録に着目し、長期にわたり高い精度を保ちながら収集する方法をわが国にいかにとり入れられるかを明らかにする目的で訪問し検討した。

結果としては多くの点でわが国の全国がん登録と同じ方法で行われ、同じような課題をもっていた。一方で多くのスタッフが携わり、知識や経験を蓄える活動は取り入れた内容であった。

A. 研究目的

2013年12月に成立し、2016年1月から施行されているがん登録等の推進に関する法律（以下「がん登録推進法」）によって、ようやく長期にわたりわが国全体のがん罹患の推移を明らかにできる仕組みが出来上がり機能するようになった。

今までは、各地域の都道府県がん登録室がいかに悉皆性の高い、高精度のがん情報

を収集するかに力を注いだが、これからは収集したデータをどのように活用し、公衆衛生や医療に生かしていくかが課題となる。

わが国においては正しくがん罹患数を数えるために、どうしても個人情報を収集しなければならない。またコホート研究をはじめとする非匿名データを必須とするものもあり、情報利用にあたっては個人情報を取り巻く様々な場面がある。そのため全国

がん登録の情報の利用と提供については多様な角度から検討され、討議されている。

一方で米国の **Surveillance, Epidemiology, and End Results Program (SEER 事業)** のようなデータが整備されれば、さらに多くの研究者や行政関係者が身近にデータ利用を進めることができ、社会に役立つことが期待される。SEER 事業にデータを提供しているユタ州がん登録を訪問し、長期にわたり高い精度を保ちながら収集する方法をわが国にいかにとり入れられるかを明らかにする目的で検討した。

B. 研究方法

2017年8月29日、ユタ州がん登録室を訪問し、担当者から直接情報収集を行い質疑や意見交換を行った。①ユタ州がん登録室の総括、制度や組織について、②実際のデータ収集から提出までの一連の流れ（ロジカルチェック、同一人物処理、集約、データ提出といった一連の **population based cancer registry** の手順）、③安全管理措置、④リンケージや変換作業、⑤研究利用 と多岐にわたり学ぶことができた。都道府県がん登録室ではユタ州がん登録室同様、データ収集、照合作業、がん情報のロジカルチェック、集約作業、データ提出、年に1回の報告書の発行として集計・統計作業を行っている。また安全管理措置についても全国がん登録を機に整備され、都道府県がん登録室で運用されてきた。よって、本報告書では実務の流れと安全管理措置を中心に結果、考察を述べる。

(倫理面への配慮)

本研究を実施するにあたり、倫理面で配慮すべきことは発生しない。

C. 研究結果

分担研究者として、筆者は実務を教育研修する立場にもあるため、②実際のデータ収集から提出までの一連の流れ③安全管理措置を中心に成果を述べる。

データ収集、研修・教育、院内がん登録 (SEER が集めている情報)

(1) データ収集の対象：1. ユタ州住民の居住する地域すべて (**population based cancer registry**) 2. 医療機関(当該地域にある)によって提出された州外の症例、3. さらなる症例登録もれを防ぐためにケースファインディングを行う(早期の前立腺癌、悪性黒色腫、慢性骨髄性白血病等)

(2) SEER の報告すべき新生物：上皮内がんと浸潤がん、悪性以外の脳・中枢神経腫瘍。(基底細胞癌と皮膚の扁平上皮癌を除く)

(3) がんの情報源：1. ユタ州の病院 48 施設、2. 放射線センター7 か所、3. 診療所、4. 病理研究所(ラボ)、5. 死亡診断書、6. 老人ホーム (Nursing Home)、7.他州とのデータ共有

(4) 病理レポート収集： e-Path を通して受け取る電子的病理報告収集(**Electronic Pathology**)と **Non Electronic pathology**(ファックス・メール・安全 FTP による)、病院の病理、病理研究所、皮膚科・泌尿器科・腫瘍内科に提携病理研究所を調査し、登録もれを確認。

これらすべての病理性情報をユタ州がん登録室で確認。

(5) 病理報告による、遡り調査

(6) 院内がん登録データ提出 (abstract)
電子データ(診断から 6 カ月以内に提出)

(7) 死亡診断書と死亡診断による遡り調査 (2014 年症例で DCO:Death certificate only : 死亡票以外の情報がないものは 0.34% : ちなみに SEER は 1.5%未満が基準)

(8) 品質管理研修

CTR (がん登録士) Certified Tumor Registrar が専門性を獲得、コーディングルールの評価や要約・届出票 abstract のチェック、院内がん登録と連携。学会や Web からの教材提供を受ける。

症例の集約・編集とデータ提出準備

症例登録の流れとして、データ受け取り→登録対象かをスクリーニングする→同一人物照合→腫瘍集約→目視 (編集確定) →品質管理チェック→提出準備完了というステップを踏む。がん登録は 1 腫瘍 1 登録が基本であり、まず照合 (同一人物か否か) を行い、ついで腫瘍を集約することによって、二重に登録しない仕組みをとっている。

患者レベルの照合、腫瘍レベルの集約には電子的な自動集約の他に目視集約が丹念になされていた。テキストデータとコードの確認やコードの精度を検証し、完全性を評価し登録士にフィードバックを行い、常に技術向上体制を保っていた。

データの編集は論理的かを確認するが、データやノウハウを積み重ね、技術や能力に磨きをかけていた。SEER、NPCR、local edits はそれぞれの患者データにおいて自動で実行するが、目視にあがるものもあり、すべてがデータ提出前に明らかにされる。

情報源としては、病院からの届出、病理

レポート、診療所などからの届出、死亡診断書、他州からの届出がある。

データ安全管理、データ構造とデータ移送

データ安全管理の目的は顕名データ (identifiable data) の保護である。Identifiable data の定義はわが国においても議論され整理されつつある。法律・ルール・協定契約が整備されていた。

安全管理必要事項は下記のとおり。

- (1) 安全管理に関する文書化
- (2) 顕名データへのアクセス権限の制限
- (3) 物理的・電子デバイス安全管理
- (4) 顕名データへのアクセスを最小限にする
- (5) データ移送や蓄積時の暗号化
- (6) スタッフへの安全管理教育
- (7) 危機管理報告(事故対応)

D. 考察

がん登録推進法によってようやくわが国のがん登録全体が精度向上することができた。今後はデータ利用にも力を入れて行かなければならない。一方でユタ州がん登録は長期にわたり、質の高いデータを SEER 事業として提供することによって公衆衛生・がん対策に尽力してきた。

わが国において、高精度のデータ収集、対策・研究への活用まで成熟した登録室は非常に少なく実行が難しいと考えていたが、実際訪問することによって、意外にもわが国の制度や仕組みが近いことを認識した。同じ目的 (population based cancer registry) であるため、当然といえるが仕組みや課題がわが国の内容と近かった。

しかし異なる点についてもあり、これら

は検証しどういった対応や対策を講じれば実行可能かも考えた。

わが国の制度と近似する点異なる点

データ収集

同じ目的とするものであるため同じであるのは当然であるが、データ収集では病理報告書等からの大量データとりこみが特徴的であった。また、幅広くデータを収集して、ユタ州がん登録室のスタッフが整理し、指導的立場である印象であった。追加項目を調査収集したりする場合、リードする存在として都道府県に登録室（全国がん登録室）の人材の養成が必須と思われた。

わが国の場合、主に国が指定するがん診療連携拠点病院およびそれに準じて都道府県が推薦する医療機関が中心に院内がん登録情報（病院の病理所見を中心とする abstract 届出）を提出することで全国がん登録に貢献している構造である。病理登録はユタ州では約 5%のもれをカバーしていたが、日米の病院機能の点からわが国で病理登録を行った場合、もれをどの程度把握できるかは不明である。

診療所における前立腺癌や、悪性黒色腫、慢性骨髄性白血病等の登録がもれることが課題とのコメントであったが、わが国においてもある程度積極的に診療所の手あげを推進することによって、同じような課題を解決できると考えられた。

他州とのデータ交換は、地域がん登録時代、特定の地域間で行われた県境のデータ登録もれを防ぐ手段と同じであった。全国がん登録の施行によってこの問題は解決された。

DCO の低さは、徹底した拾い出しによる

ものであり、低い DCN : Death certificate notifications（死亡票により初めてがんを把握）に基づくと考えられる。わが国は 2016 年診断症例からすべての病院と指定を受けた診療所によって登録されるため、2018 年末に集計される結果と比較すべきであり、今までにない全国的に低い DCN・DCO が期待される。

研修・教育

わが国では、院内がん登録において初級・中級の試験が開始され管理更新も厳しく行われ始めた。モチベーションを保ちながら精度を保ち続けるためには、最終的に人材育成に力をいれなければ継続しない。

わが国においても、学会やがん診療連携協議会がん登録部会など相互に高めあえる環境を整備し継続させなければならない。

集約・データ提出準備

同一人物処理は日本独自で非常に困難な課題と考えていたが同様の課題がユタ州がん登録室にもあることに驚いた。また、それらの一つずつ解決していくことによってデータが結果的に成熟することを再認識した。また、自動集約についても実施されているものの、最終的な判断は知識と経験が必要であった。

わが国においても自動集約によって実務者の仕事量は減少したが、質を落とさないためにもさらなる訓練を引き続き行わなければならない。

安全管理

わが国の安全管理措置マニュアルが必要項目を網羅していたことを確認した。がん登録推進法によって、物理的・技術的安全管理はかなり統一され、実務側の負担が軽減したが、スタッフへの安全管理教育など

が今後の課題となる。また、都道府県におけるがん登録審議会での対応が迫られるため、個人情報の **identifiable data** についてさらなる議論や定義の確認が必要である。

その他(マンパワー)

ユタ州がん登録室を訪問して、わが国のがん登録推進法施行後において、SEER 事業のようなデータを作成することは技術的に実現不可能ではないが、継続するためのスタッフや財源についても考慮していかなければならない。

ユタ州は筆者の管轄する栃木県の 1.5 倍程度の人口だが、関わっているスタッフの数は多く (すべてが専従ではないが、27 名がホームページで紹介されている¹⁾) 分担内容も多様であった。

E. 結論

大きく進んでいると考えていた SEER 事業であったが、わが国の制度でも類似点が多く技術的に可能かもしれない。しかし仕組みとして継続するためには人材育成・マンパワーの充実といった長期的な視点から総合的に取り組まなければならない。

(参考文献)

1) <http://uofuhealth.utah.edu/utah-cancer-registry/about/our-team.php> (2018/3/31)

F. 健康危険情報

(総括研究報告書にまとめる)

G. 研究発表

1. 論文発表

該当なし

2. 学会発表

1. 大木いずみ, 西野善一, 松田智大. 地域

がん登録データを用いたがん診療実態の把握. 第 76 回日本公衆衛生学会総会, 鹿児島, 2017

2. 渡辺良光, 福田芳彦, 富田倫子, 大木いずみ, 他. がん登録を用いたがん検診精度管理支援. 第 55 回栃木県公衆衛生学会, 栃木, 2017

3. Okii I, Nishino Y, Saruki N. Profile of the Japanese Association of Cancer Registries. Asian Cancer Registry Forum 2018, Bangkok, Thailand, 19-21 March, 2018.

H. 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

全国がん登録情報を基盤とした長期記述疫学研究用の特定匿名化情報の整備について
-米国 SEER にデータを提供しているユタがん登録訪問により見えてくる課題-

研究分担者 伊藤秀美 愛知県がんセンター研究所 遺伝子医療研究部 室長

研究要旨

2016年1月から施行された全国がん登録等の推進に関する法律のもとに収集される全国がん登録情報を基盤とした長期記述疫学研究用の特定匿名化情報の整備のため、米国の人口単位のがん登録である surveillance, Epidemiology, and End Results Program（以後 SEER）のデータ収集、データ提供のしくみを評価し、SEER へデータを提供するユタがん登録室での、データ収集や追跡、照合、シェアの現状を聞き取り調査することにより、全国がん登録情報を基盤とした長期記述疫学研究用の特定匿名化情報の収集や提供において、米国の SEER のしくみをモデルとした場合の実現可能性や課題などを考察した。

A. 研究目的

2016年1月から施行されているがん登録等の推進に関する法律第3条（基本理念）第3項では、がん対策の充実のためには、全国がん登録の実施のほか、がんの診療の状況を適確に把握することが必要であることに鑑み、院内がん登録により得られる情報その他のがん診療情報の収集が図られなければならないとされているが、がん診療情報の収集の方法は具体的に示されていない。

本研究では、そのひとつの可能性として、米国の人口単位のがん登録である SEER のデータ収集、データ提供のしくみを学び、全国がん登録情報を基盤とし、長期にわたり院内がん登録やその他の既存情報とリンケージするしくみ、「特定匿名化情報」の提供のしくみ（日本版 SEER）の整備のための課題を洗い出した。

B. 研究方法

SEER のしくみを理解するため、1) SEER やデータを提供する州がん登録のひとつであるユタがん登録の公開情報を精査した。2) ユタがん登録室を訪問し、聞き取り調査を実施した。訪問するに辺り、HP など公表されている情報を収集した。

（倫理面への配慮）

本研究を実施するにあたり、倫理面で配慮すべきことは発生しない。

C. 研究結果

1) 公開情報の精査

<http://healthsciences.utah.edu/utah-cancer-registry/>

上記の HP によると、ユタがん登録では、クリニックなどから Utah 州法 (R384-100 Cancer Reporting Rule) で定められている項目を、また州の協力病院から、SEER や北米中央がん登録協議会 (NAACCR) のデータ規格にあった詳細な診断、治療、予後情報を、顕名データとして収集していることが分かったが、具体的な収集方法や、匿名化のタイミングについての情報を得ることはできなかった。

2) 聞き取り調査

ユタがん登録を訪問し、下記について聞き取り調査を行った。

- ・データの基準について
- ・データ収集について
- ・データ編集、集約方法について
- ・データの機密保持について
- ・追跡について
- ・データ照合
- ・データシェア

このうち、特に、データのリンケージならびにデータシェアについての聞き取り調査の結果を示す。

①届出症例の追跡について

ユタがん登録では、積極的ならびに消極的な追跡調査を実施している。情報源として、社会保障庁（70%）、院内がん登録（15%）高齢者向け、低所得者向けの公的医療保険サービスセンター（6%）、その他、州の死亡診断書、担当医師への問い合わせ、患者への連絡、有権者登録ファイルな

どであった。国の死亡情報と社会保障庁データとの照合については SEER のサポートの元、実施されている。また、ほとんどが電子的なリンケージによるものである。追跡率は年齢により違うが、0-64 才で 90%、65 才以上 95%であるが、上皮内がんでは 90%と追跡率はやや劣っていた。ユタ州における追跡率は、99%と高いが、小児における追跡率は低いのが(90%以上)、課題である。

②照合とデータ交換のプロセスについて

データ照合の目的は、個人照合、腫瘍データ集約、追跡調査：最終生存確認日の更新、コホート研究におけるがん罹患把握であった。主なリンケージ方法は、一致の可能性をスコア化する確率的な照合方で、個人識別指標は、姓、名、生年月日、社会保障番号で、住所は照合の確認に利用されている。社会保障庁や公的医療保険サービスセンターでのデータ照合には社会保障番号が必要である。また、社会保障庁、公的医療保険サービスセンター、国の死亡情報や州の有権者登録ファイルを使った消極的な追跡は、主に生死を確認するために実施するが、その他、インディアン保健サービスから人種情報、Oncotype 診断検査からはバイオマーカーや腫瘍マーカー、健康保険請求データベースから治療情報を得るために、追跡調査を実施していた。また、ユタがん登録では、疫学研究実施のため、コホート研究データとの照合も請け負っていた。照合ソフトは、SEER の開発した SEER Link Plus を利用していた。

データ交換に関して、州間でデータ交換をしてダブルカウントを防ぐようなくみもあり、ユタ州においても近隣州とデータ交換をしていた。また、NAACCR によるデータ交換協定により、米国全州と電子的にデータ交換をした上で、SEER へ匿名化データを提出しているため、州間の重複は無視でき、さらにユタ州では近隣州との移動も少ないため、重複は無視できる程度であるとのことであった。また、コホート研究におけるがん罹患把握のための照合の効率を上げるため、National Cancer Institute が Virtual pooled registry システムを開発中であるが、このしくみにより、参加者が複数の州にわたるコホート研究の場合、研究者が参加者のがん罹患を把握するのに、各州の倫理審査委員会で承認を得ることなく中央の倫理審査委員会の承認のみで照合が可能となる。

③ユタがん登録を使った研究とがんサーベイラ

ンスについてデータ開示とその条件

ユタがん登録情報の提供において、ユタ保健局はデータの機密保護と守秘義務、データ開示状況に関する覚書を結んだ上で提供している。また SEER や CDC へのデータ提供は、データの機密保護および被験者保護の連邦ガイドラインと、がん研究促進のため要件を含む契約を結んだ上で提供している。ユタ大学へのデータ提供においては、データの機密保護と守秘義務指針を定めている。その他、研究に対する諮問委員会でユタがん登録情報の提供に関して審議されている。

ユタがん登録では、データセット作成、診療録抽出、研究参加者へのコンタクト、研究資源(研究計画を立てるための基礎データ提供、データリンケージ、Utah Population Database への提供)など、がん研究を支援していて、ユタがん登録データを使った、査読のある論文は 2011-2016 年で 287 (2016 年 7 月から 1 年で 35) であった。

D. 考察

全国がん登録情報を基盤とした長期記述疫学研究用の特定匿名化情報(日本版 SEER の整備のため、米国の SEER における情報収集と提供のしくみについて、特に匿名化情報を SEER に提供する地域のがん登録室の視点で、調査し、日本版 SEER の整備における課題点などを考察する。

全国がん登録のデータ収集、データ編集、集約方法、追跡、データ照合という観点において、ユタがん登録と大きな差異はなかった。それにも関わらず、SEER プログラムにおいて、住民単位の高精度な詳細情報を長期間計扱い属的に提供できている理由のひとつとして、ユタがん登録など SEER プログラムに参加する州が、SEER から継続的な金銭的、技術的支援を受けていることが考えられる。実際に、ユタがん登録室では、SEER への匿名化情報の提供に関する契約を SEER と結び、十分な資金の提供を受けていた。この SEER からの支援により、がん登録室は高精度な詳細情報を継続的に SEER へ提供することが可能となっている。参加登録に対して、持続的な金銭的サポートを提供できる仕組みを作ることが、日本版 SEER の整備には必要であろう。

F. 健康危険情報

(総括研究報告書にまとめて記入)

G. 研究発表

1. 論文発表

1: Improvement in 5-Year Relative Survival in Cancer of the Corpus Uteri From 1993-2000 to 2001-2006 in Japan. Inoue S, Hosono S, Ito H, Oze I, Nishino Y, Hattori M, Matsuda T, Miyashiro I, Nakayama T, Mizuno M, Matsuo K, Kato K, Tanaka H, Ito Y; J-CANSIS Research Group. J Epidemiol. 2018 Feb 5;28(2):75-80.

2: Potential overtreatment among men aged 80 years and older with localized prostate cancer in Japan.

Masaoka H, Ito H, Yokomizo A, Eto M, Matsuo K. Cancer Sci. 2017 Aug;108(8):1673-1680.

3: Recent Improvement in the Long-term Survival of Breast Cancer Patients by Age and Stage in Japan. Yoshimura A, Ito H, Nishino Y, Hattori M, Matsuda T, Miyashiro I, Nakayama T, Iwata H, Matsuo K, Tanaka H, Ito Y. J Epidemiol. 2018(in press)

2. 学会発表

1: Prognostic impact of tumor location in colon cancer: the Monitoring of Cancer Incidence in Japan (MCIJ) Project. Nakagawa H, Matsuda T, Ito H. 39th IAACR Scientific Conference. Utrecht, the Netherlands, Oct, 2017 (Poster Presentation)

2: 中川弘子、伊藤秀美ら. 愛知県がん登録の照会. 第26回全国がん登録協議会学術大会. 2017年6月9-10日(愛媛、ポスター発表)

3: 山口通代、伊藤秀美ら. 愛知県における国・県拠点病院の診療実態の把握と5年相対生存率改善度の試算. 第26回全国がん登録協議会学術大会. 2017年6月9-10日(愛媛、ポスター発表)

H. 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

特記すべきこと

厚生労働科学研究費補助金（がん対策推進総合研究事業）
（分担）研究報告書

全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究
- 記述疫学研究担当 -

研究分担者 林 櫻松 愛知医科大学医学部公衆衛生学 准教授

研究要旨 研究目的で全国がん登録データを使用する際の参考となる資料を得るため、SEER Data の利用申請方法、通常利用申請でどの程度の記述疫学ができるかを検討した。さらに、日米とも増加傾向にある膵がんについて、SEER*Stat を用いて、罹患率・死亡率の年次推移や、診断時ステージ、組織型の分布、ステージ別 5 年生存率に関するデータ解析を実施した。通常利用申請で入手した SEER Data には日系アメリカ人のがん統計データが含まれていないため、別途に Asian Pacific Islander Specialized Population Dataset の利用申請を行った。2000 年以降、米国では膵がんの罹患率も死亡率も上昇傾向が続いている。診断時ステージ分布や 5 年生存率についての集計結果から、依然として早期発見が困難で、予後が大変不良であることが分かった。日本側の膵がんデータとして精度の高い、3 県の（山形、福井、長崎）がん登録データの利用申請を行い、膵がんの罹患率や死亡率の日米比較を行う予定である。

A. 研究目的

米国 National Cancer Institute の Division of Cancer Control and Population Science が運営している SEER Program は、最大で 18 州からがん登録・診療情報の提供を受け、全米人口の約 28% をカバーしている。SEER Data を利用したがん記述疫学研究は幅広く行われている。全国がん登録データを使用する際の参考資料を得るため、SEER Data の利用申請方法や通常利用申請でどの程度の記述疫学ができるかを検討した。さらに、日米とも増加傾向にある膵がんについては、SEER Data の解析ソフトである SEER*Stat を用

いて、罹患率・死亡率や、診断時ステージ、組織型の分布、ステージ別 5 年生存率に関するデータ解析を実施し、日本の地域がん登録データと国際比較を行う。

B. 研究方法

SEER Data の利用申請が許可後、SEER ウェブサイトから SEER*Stat をダウンロードした。Client-Server Mode により、SEER*Stat サーバーにアクセスし、SEER Data の解析を行った。Client-Server Mode では、インターネット接続形式で SEER 19 州のがん登録データにアクセスできるほか、死亡率や遅延補正がん罹患データの解析も可

能である。膵がんの罹患率、死亡率などについての日米比較を行うため、別途に Asian Pacific Islander Specialized Population Dataset に含まれている日系アメリカ人の膵がん罹患データも入手した。

米国での膵がん罹患率・死亡率の年次推移や診療情報を検討する際に用いられた SEER データセットは、SEER 8 Regs Research Data, Nov 2016 Sub(1973-2014)<Katrina/Rita Population Adjustment>、SEER 9, plus remainder of CA and NJ, Nov 2016 Sub (1990-2014) detailed API plus White Non-Hispanic-populations、released May 2017, based on the November 2016 submission、US mortality data, 1969-2014(NCHS) である。Joinpoint analysis により罹患率や死亡率の経年的変化を検討した。

(倫理面での配慮)

本研究は、利用申請が SEER に承認されたもので、匿名化が済んだ集団データで統計解析を行った。

C. 研究結果

1) SEER Data の利用申請とデータの内容について

SEER*Stat を用いて、がん種別、性別、組織型別、年齢階級別の基本集計のほか、がん診断時ステージやステージによる 5 年生存率などの分析も可能である。一方、社会経済状況や保険情報なども含まれる SEER-Medicare Linkage Data の利用申請をしたものの、米国国外からの研究者が申請できないことが判明した。日系

アメリカ人における膵がん罹患率や死亡率のデータは、Asian Pacific Islander Specialized Population Dataset に含まれており、別途に利用申請が必要であるため、研究の目的や意義、解析予定項目、データ発表などを含めた 1 枚の申請文を提出し、承認を受けた。

2) 膵がん記述疫学の日米比較について

SEER*Stat で米国での膵がん罹患率・死亡率、生存率、臨床診療情報についての検討を行った。遅延補正罹患率(Delay-adjusted incidence)は、1974 年から 2014 年の期間中に 10 万人あたり 11 から 13 あたりで推移しており、特に 2000 年以降は緩やかな上昇傾向を続けている。2001 年から 2008 年の平均年変化率(APC)は 1.77 と統計学的有意であった(図 1)。死亡率については、2001 年以降は上昇傾向を続けており、2001 年から 2008 年までの期間中に APC は 1.77、2008 年から 2014 年までの期間中に 0.26 といずれも統計学的に有意であった。

2007 年から 2013 年に SEER 9 地域で報告された 60171 症例を対象として、診断時ステージの分布を検討した結果、遠隔転移が 52%と最も多く、ついでに領域(25%)、限局(10%)の順であることが分かった。

生存率解析では、遠隔転移が認められた膵がん患者の 5 年生存率は 2.7%と著しく予後が不良であった。一方、限局がんに残っている場合、5 年生存率は 31.4%であった。2010 年から 2014 年までに報告された膵がん症例 40403 人の組織型分布では、腺癌が 81.1%を占めており、最も多いタイプとなっている。

日系アメリカ人における膵がんの罹患率

の年次推移では、1990年から2014年までの期間中に上昇傾向が認められ、APCは1.49であった。特にハワイ在住の日系アメリカ人に限った分析では、APCは男性で3.40、女性で0.18と増加傾向が顕著であった。

日本側の膵がんデータとして、精度の高い、3県の（山形、福井、長崎）がん登録データの利用申請を行い、来年度は膵がんの罹患率や生存率などについての日米比較を行う予定である。

D. 考察

今回の研究では、SEERデータの利用申請方法、通常利用申請でどの程度の記述疫学ができるかを検討した。申請が承認後、ダウンロードしたSEER*Statを用いて、すぐにデータ解析できたという点で大変便利なものである。研究者のオフィス環境によって、SEER Dataの利用は、SEER*Statサーバーに接続した上での利用と、圧縮されたファイルをダウンロードした上での利用という2つのオプションがあることにも良さを感じた。SEER*Statを用いて、複数のデータベースのうち、自分の研究目的に合致するデータベースを選択し、がん種別、性別、組織型別、年齢階級別の基本集計のほか、がん診断時ステージやステージによる5年生存率などの分析も可能である。

本研究の主な目的である膵がんの日米比較を行うため、まず、SEERがカバーする地域における膵がん罹患率や死亡率の動向、診断時ステージ分布、予後などの解析を行った。その結果、膵がんは特に2000年以降、罹患率も死亡率も上昇傾向が続いていることが明らかであった。現在の増加傾向を続

けていけば、2030年には膵がんは死亡率の2位になるとの予測が論文発表されている。診断時ステージ分布と5年生存率を集計した結果から、膵がんは依然として早期発見が困難で、予後が大変不良であることが分かった。

本研究のもう一つの目的は、日系アメリカ人と日本人の膵がん統計データを比較することである。日系アメリカ人における膵がんの罹患率は、SEER全体のそれより明らかな上昇傾向が認められた。図2に示すように、ハワイ日系アメリカ人での膵がん罹患率の変動が大きいことが見られたため、5年間における罹患率の平均を求めるなど対応が必要である。

日本とアメリカでは、膵がんに関する記述疫学はそれぞれ実施されているものの、両国を比較した記述疫学は見当たらない。特に日系アメリカ人では罹患率の上昇は顕著であることが今回の分析で明らかであった。上昇の原因や人種差を説明する要因については、今後の研究でさらに検討する必要がある。

E. 結論

SEER Dataを用いて、膵がんの罹患率や死亡率の年次推移、診断時ステージ分布や予後状況を検討した。日系アメリカ人における膵がん罹患率の上昇傾向はSEER全体のそれより顕著であった。

F 健康危険情報

なし

G 研究発表

1. 論文発表

- 1) Nakayama Y, Lin Y, Hongo M, Hidaka H, Kikuchi S. Helicobacter pylori infection and its related factors in junior high school students in Nagano Prefecture, Japan. Helicobacter 2017;22:e12363
- 2) Tsuda M, Asaka M, Kato M, Matsushima R, Fujimori K, Akino K, Kikuchi S, Lin Y, Sakamoto N. Effect on Helicobacter pylori eradication therapy against gastric cancer in Japan. Helicobacter. 2017;22: 12415.
- 3) Wang C, Nishiyama T, Kikuchi S, Inoue M, Sawada N, Tsugane S, Lin Y. Changing trends in the prevalence of H. pylori infection in Japan (1908–2003): a systematic review and meta-regression analysis of 170,752 individuals. Sci Rep 2017;7(1):15491.

- 4) Osaki T, Zaman C, Yonezawa H, Lin Y, Okuda M, Nozaki E, Hojo F, Kurata S, Hanawa T, Kikuchi S, Kamiya S. Influence of Intestinal Indigenous Microbiota on Intrafamilial Infection by Helicobacter pylori in Japan. Front Immunol 2018;9:287

2. 学会発表

中柄 昌弘、林 櫻松、伊藤 秀美、木下 文恵、小林 由美子、石井 浩、笹平 直樹、上野 誠、江川 直人、来間 佐和子、中尾 春壽、松田 文彦、菊地 正悟、松尾 恵太郎：一般日本人集団における膵がん発症予測モデルの構築、日本人類遺伝学会第 62 回大会 2017 年神戸

H. 知的財産権の出願・登録状況
特に無し

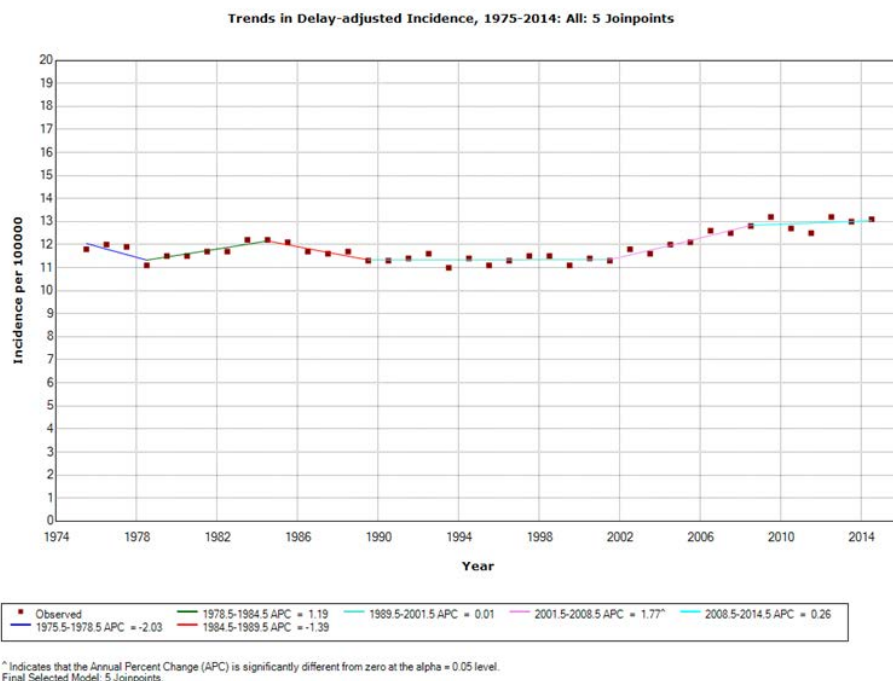
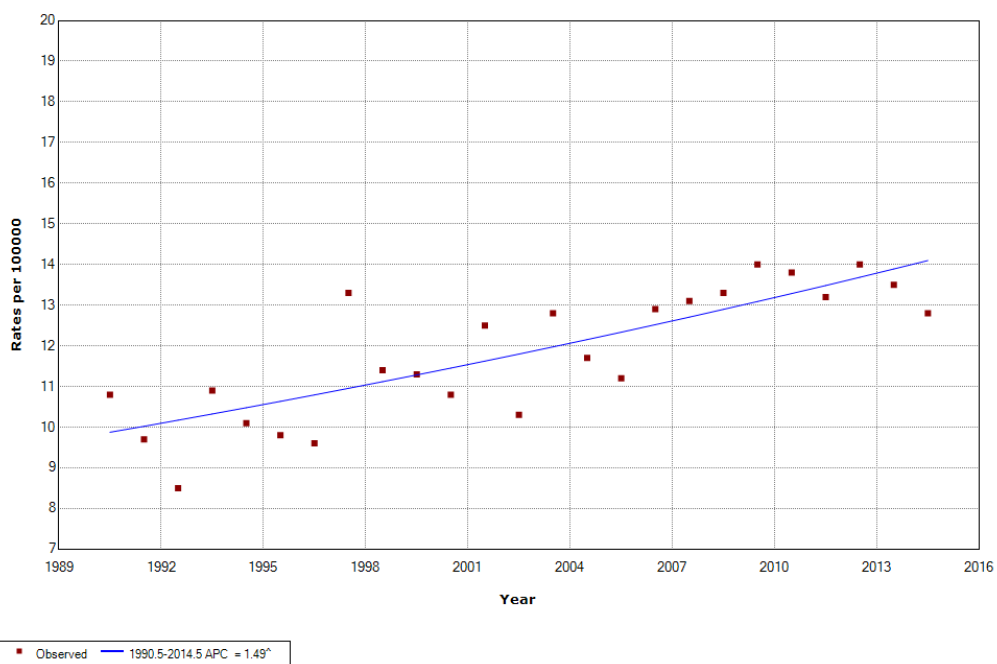


図 1 Trends in Delay-adjusted Incidence, 1975–2014: All: 5 Joinpoints

Trend in age-adjusted incidence rates for Japanese American 1990-2014: All: 0 Joinpoints



^ Indicates that the Annual Percent Change (APC) is significantly different from zero at the alpha = 0.05 level.
Final Selected Model: 0 Joinpoints.

☒ 2 Trends in age-adjusted incidence rates for Japanese American, 1990-2014: All: 0 Joinpoints

厚生労働科学研究費補助金（がん対策推進総合研究事業）

（分担）研究報告書

全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究
- 記述疫学研究（国際比較）疫学研究として活用するためのデータベースの条件 -

研究分担者 井上真奈美 国立がん研究センター社会と健康研究センターコホート連携研究部 部長

研究要旨

米国NCIとユタ州のSEERプログラムを視察した。特に研究者に提供しているSEERデータベースの構成について確認し、日本版SEERへの適用性と、評価に耐えうる質の高い疫学研究に利用するための条件を考察した。その結果、米国SEERデータベースと同様に、日本版SEERについては、他のデータベースとのリンケージを想定せず、単独利用を前提とした提供をめざすのが現実的かつ効率的であると考えられた。その前提で、人口ベースであるという特性を維持しながら、がん診断・治療・予後の長期の経年推移（トレンド）を捉えられる項目を網羅すること、これを類似のデータとプールし個票単位のデータとして統合解析を実施できるようなフォーマットに落とし込んでおくのが今後の活用推進に向けた戦略として重要である。

A. 研究目的

米国ではすでに National Cancer Institute (NCI) によって、40 年間にわたり The Surveillance, Epidemiology, and End Results Program (SEER 事業) が実施されている。この事業では協力州からがん診療情報を含む地域がん登録情報の提供を受けて SEER Data として整備し、簡単な申請で研究者に提供しているため、標準公表集計値以外の部位や組織型の罹患率や生存率の記述疫学研究等に幅広く活用されている。わが国のがん登録推進法では、法律第 3 条の基本理念において、その他のがん診療情報の収集と活用を掲げ、2019 年以降、2016 年罹患の全国がん登録情報の提供が予定されているが、その具体的方法はまだ決まっていない。

本分担研究では、米国 SEER プログラムを訪

問して、そのデータ提供の枠組みや形体について視察を行い、特にがん疫学研究の観点から、ヘルスポリシーに貢献できるエビデンス創出のための日本版 SEER データのあり方について考察した。

B. 研究方法

米国 NCI とユタ州の SEER プログラムを視察した。特に研究者に提供している SEER データベースの構成について確認し、日本版 SEER への適用性と、評価に耐えうる質の高い疫学研究に利用するための条件を考察した。

(倫理面への配慮)

本報告に関しては、データベースを利用した研究を実施していない。

C. 結果と考察

米国版 SEER データは、外部研究者の利用に際して、外部データベースとのリンケージは想定せず、SEER データのみの解析を前提としている。全国がん登録情報とのリンケージによる研究は、日本版 SEER データとは異なる全国がん登録情報本体のデータ利用の枠組の中で整備が進められており、そのため、日本版 SEER については、他のデータベースとのリンケージを想定せず、単独利用を前提とした提供をめざすのが現実的かつ効率的である。したがって、以降は単独データを提供するという前提とする。

疫学研究からみた SEER データの強みは、人口ベースであること、長期の経年推移（トレンド）が見られること、がん診断・治療・予後に関する情報があることである。これらの特性を維持しながら、ヘルスポリシーに通じるエビデンスを創出することが、日本版 SEER データの有効活用のキーになる。

人口ベースであることは、日本版 SEER のように人口ベースがん登録の提供するデータベースでは必要不可欠な要素である。もしこの要件が欠ければ、類似の病院施設などで任意に作成されている患者登録データベースでも代用できてしまい、日本版 SEER をあえて作成する必要性もなくなってしまう。人口ベースでの検討は単に治療などの生存率への影響を評価できるのみでなく、その集団（国や地域）へのインパクトを評価できるという利点がある。

長期の経年推移（トレンド）が見られることとは、複数年（できるだけ長期）の診断について情報の含まれていることである。全国がん登録を対象とするなら、2016 年からのみ可能となるため、しばらくの間、この要件を満たすことはできない。しかしながら、対策のためのエビデンスが求められている現在、長期情報が網羅できないために解析が実現できないのは方

策としてはふさわしくない。そのため、1975 年から 2015 年までについては、これまでの全国推計値を代用することが現実的であり、検討すべきである。

進展度や治療、生存など、がん診断・初期治療・予後に関する情報があることについては、罹患数のカウントのみではなく、治療や診断法の効果について検討が人口ベースで可能になるための不可欠な要素である。一方、がんの進行度や治療内容など、コアとなる収集項目以外の詳細情報について、必ずしも全国がん登録の中で十分に収集できていない情報も存在し、それらが、日本版 SEER データを用いた治療効果や生存率分析の中心的役割を担うことも事実である。十分な収集ができていない項目については、院内がん登録など他の情報源との連携により、情報を補完していくのが不可欠である。いずれにしても、人口ベースという条件は保持することが肝腎であり、例えば情報充足率の高いがん登録に限定する場合も、人口ベースを崩さないようにする努力が必要である。

最後に、記述疫学の国際比較を考えた場合、単なる集計結果の比較ではなく、日本版 SEER データベースの個票単位データを米国 SEER や他国の同様のデータベースと統合して行うプール解析は、近年ますます研究ニーズが高まっている利用の一つである。この方法では、結果に影響を与える交絡因子の影響を共通にコントロールすることが可能であるため、そうでない研究と比較して、相対的に研究の質を高く保つことができる。既に International Association of Cancer Registries (IACR) で実施されている CONCORD(参考文献)などの地球規模比較研究との線引きが必要であるが、少なくとも全世界ではなく、2~3 国間の治療予後や民族差などを検討する際には、有効かつ高い質を維持できる研究方法として実施を促進すべきと考える。そのためには、共通項目につい

では、例えば米国 SEER データなどの既存の同種データベースと書式や項目カテゴリーをできる限り統一して、多変量解析、メタ解析に耐えうるフォーマットに今から仕上げておくのが肝要である。

D. 結論

米国 SEER データベースと同様に、日本版 SEER については、他のデータベースとのリンケージを想定せず、単独利用を前提とした提供をめざすのが現実的かつ効率的である。その前提で、人口ベースであるという特性を維持しながら、がん診断・治療・予後の長期の経年推移（トレンド）を捉えられる項目を網羅すること、これを類似のデータとプールし個票単位のデータとして統合解析を実施できるようなフォーマットに落とし込んでおくのが今後の活用推進に向けた戦略として重要である。

参考文献

Allemani C, Matsuda T, et al, CONCORD Working Group. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis

of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. 2018 Mar 17;391(10125):1023-1075.

E. 研究発表

1. 論文発表

なし

2. 学会発表

なし

F. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

厚生労働科学研究費補助金（がん対策推進総合研究事業）
（分担）研究報告書

全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究
- 全国がん登録情報利用者に求められる安全管理措置 -

研究分担者 西野善一 金沢医科大学医学部公衆衛生学 教授

研究要旨

全国がん登録利用者に求められる安全管理措置を検討する一環として、第10回厚生科学審議会がん登録部会で提示された「全国がん登録 利用者の安全管理措置（仮称）（案）」が持つ課題を検討した。特に、非匿名化情報の保管場所の限定、利用場所に求められる環境（物理的安全管理対策）の明確化、取得時の安全管理対策、に関する記載が課題として考えられた。

A. 研究目的

「がん登録等の推進に関する法律」では、第33条で全国がん登録情報もしくは都道府県がん情報の提供を受けた者に対しても秘密保持義務を課すとともに、同法第6章で秘密を漏らした者への罰則が規定されている。

本研究では、全国がん登録情報利用者に求められる安全管理措置を検討する目的で以下を実施した。

B. 研究方法

平成30年3月15日の第10回厚生科学審議会がん登録部会における資料として提出された「全国がん登録 利用者の安全管理措置（仮称）（案）」（以下「利用者安全管理措置」）の内容が持つ課題を検討した。

C. 研究結果

「利用者安全管理措置」の内容につき今後さらなる検討が必要と考えられる事項は

以下のとおりである。

1. 一般的事項

（1）非匿名化情報の保管場所の限定

「利用者安全管理措置」では複数の利用場所における情報の保管を想定しているが、特に非匿名化情報については保管場所を最低限とすることを求める必要がある。全国がん登録において非匿名化情報を利用する主な目的は利用者が持つ研究データとの突合によるデータへの全国がん登録情報の付加であると想定される。その際、非匿名化情報の保管場所は、突合が行われる統括利用責任者の施設に限定し、その他の利用場所には統括利用責任者がデータの匿名化を実施して利用（責任）者に送付することが望ましい。合わせて統括利用責任者から利用者に送付する解析用データにつき匿名化の基準を示す必要がある。

（2）利用場所に求められる環境（物理的安全管理対策）の明確化

「利用者安全管理措置」では利用場所が独立していないことも想定して安全管理措置を定めているが、非匿名化情報の保管場所については独立した部屋（利用者の個室、専用の作業室）での保管を原則とすることについて検討する必要がある。その上で、利用場所が独立した部屋である場合と共用である場合のそれぞれについて、必要な技術的安全管理対策を定めることが望ましい。

（3）取得時の安全管理対策の記載

窓口組織から情報を受領した後のデータの処理に関する安全管理対策を「利用者安全管理措置」に記載する必要があると考える。特に提供に使用された電子媒体の処理方法について規定する必要がある。盗難、紛失等のリスクを考慮すると、バックアップデータを窓口組織が保有することを前提に、提供データを含む電子媒体は PC 等へのデータの移行後は廃棄するよう定めることが望ましいと考える。

2. 具体的事項

（1）組織的安全管理対策

1) 「レセプト情報・特定健診等情報の提供に関するガイドライン」に記載されているような運用管理規程等を総括利用責任者が定めることについて検討を要する。その場合、運用管理規程等には、対策（2）で定める利用者のリストや、対策（6）で定める事故時対応手順を含めるものとし、運用管理規程の例を「利用者安全管理措置」に示すことが望ましい

（2）物理的安全管理対策

1) 先に述べたように利用場所に求められる環境（設置要件）をより明確にする必要がある。

2) 対策（2）は「複数の鍵を更に鍵付きボックスに収納して、」とあるが、利用にあたって全国がん登録に関する情報を含んだ電子媒体及び紙媒体を複数のキャビネットに保管する必要性が想定し難い。原則として1つのキャビネットにまとめて保管されるべきという観点から「複数の」という表現は検討を要する。

3) 対策（12）の二重の鍵による施錠を求める点については、都道府県がん登録室に対する安全管理対策でも必須としていないことから整合性の点で疑問がある。また、本対策は利用者自らが対応できる範囲を超えている可能性があり、利用にあたっての妨げとなる恐れがある。

（3）技術的安全管理対策

1) 対策（4）で個人情報を取り扱う PC およびサーバを生体計測+ID・パスワード等の 2 要素認証としているが、生体計測による認証の整備は利用者の負担となる可能性があり検討を要する。

2) 対策（5）のパスワードはより具体的な要件を提示することが望ましい。例えば、英字（大文字と小文字を含む）、数字、記号を含む 10 桁以上とすることなどが考えられる。

3) パスワードが記録された用紙、ファイルは個人情報と同様の安全管理措置にて保管することを明記する必要がある。

4) 「レセプト情報・特定健診等情報の提供に関するガイドライン」では、情報を保管するシステムのアクセスの記録とアクセスログの確認を求め、アクセス機能がない場合は業務日誌等で操作の記録（操作者及び操作内容）を必ず行うこととしている。また、アクセスログへのアクセス制限および

不当な削除/改ざん/追加等を防止する対策等を講じることとしている。「利用者安全管理措置」においても利用場所の環境に応じてこれらの必要性を検討する必要がある。

5) 利用場所の環境に応じてクリアスクリーン等の対策が必要である。

(4) 移送

1) 利用者による非匿名化情報の分散保管を防ぐために、統括利用責任者からの個人情報移送は原則として禁止することなどを検討する必要がある。

2) 対策(5)で定める「強固な暗号化方法」につき具体的に示すことが望ましい。

(5) PC管理

1) 対策(4)の「管理者用パスワード」は「利用者安全管理措置」では定義されておらず修正を要する。

2) 対策(5)で「PC等のユーザ登録は、利用者が実施する」とあるが、統括利用責任者または利用責任者による実施に改める必要がある。

3) 対策(6)で規定されているユーザIDとその利用者の紐付け確認作業につき実施者(統括利用責任者または利用責任者)を明記する必要がある。

D. 考察

「利用者安全管理措置」で規定されている安全管理措置は、「全国がん登録における個人情報保護のための安全管理措置マニュアル」において都道府県がん登録室に対して必須としている基本対策の内容をふまえて、基本的に同じ水準の対策を求めている。利用者が提供を受けた非匿名化情報は都道府県がん登録室が扱う個人情報と同様のものであることを考えると、非匿名化情報の

利用者に対して基本的に都道府県がん登録室と同水準の安全管理措置を求めることは適切と考える。その上で、利用に特化して想定されるリスクをさらに分析し対策を提示する必要がある。

E. 結論

「利用者安全管理措置」では利用に特化して想定されるリスクと対策についてより具体的な記載が求められる。その際の安全管理措置は、情報の漏洩等の事故が生ずることがないように対策が求められることが当然であるが、一方で過度の規制により利用が妨げられることがないように配慮されるべきである。

参考資料

1) 厚生労働省：全国がん登録 利用者の安全管理措置(仮称)(案)

<http://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000197941.pdf>

2) 厚生労働省：レセプト情報・特定健診等情報の提供に関するガイドライン

<http://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000135460.pdf>

F. 健康危険情報

(総括研究報告書にまとめる)

G. 研究発表

1. 論文発表
該当なし

2. 学会発表
なし

H. 知的財産権の出願・登録状況

なし

1. 特許取得

3. その他

なし

なし

2. 実用新案登録