

2018年1月16日会議資料  
日本版SEER研究班  
平成29年度研究進捗確認

日本版SEER構想資料

研究の概要  
(申請時)

WE ARE A PREMIER SOURCE  
FOR CANCER STATISTICS  
IN THE UNITED STATES.



- 日本にSEERを
  - SEER
  - 米国、NCI Program、40年以上（1973～）
    - Surveillance
      - 協力州（最大18）の州がん登録情報による罹患率
      - その他のがん診療情報を時代ニーズで登録
    - Epidemiology, and
      - 疫学
        - » Time (trend), Place, Person(race)
    - End Results
      - 信頼性の高い罹患情報に基づく生存率

## 研究の目的、必要性 (申請時)

- 日本にSEERを
  - 高い精度、長期持続
    - 全国がん登録を基盤
      - 「がん診療情報」の収集の仕組み
        - » 追加収集
        - » 既存情報とのリンケージ
      - 「特定匿名化情報」の提供の仕組み
        - » 本人同意なく登録、登録情報の二次利用
        - » 信頼される匿名化

第3条（基本理念）  
第4項で努力目標

第21条（その他の提供）  
第5項、6項 提供要望の  
高いと見込まれる情報につ  
いて、あらかじめ匿名化  
（特定匿名化）して全国  
がん登録データベースに記  
録できる

## 期待される成果 (申請時)

- 基本理念第3条第3項「がん対策の充実のためには、全国がん登録の実施のほか、院内がん登録により得られる情報その他のがんの診療に関する詳細な情報の収集が図られなければならない。」に対して、未だ確立していないがん診療情報の収集のあり方を、米国で確立している仕組みから学び、我が国の法令等の実情に即して検討することで、実現可能性の高い仕組みの提案が期待される。
- 同条第4項「全国がん登録及びがん診療情報の収集により得られた情報については、これらががん患者の診療等を通じて得られる貴重な情報であることに鑑み、民間によりものを含めがんにかかる調査研究のために十分に活用されるとともに、その成果ががん患者及びその家族をはじめとする国民に還元されなければならない。」に対して、提供需要が高いと考えられる「特定匿名化情報」をあらかじめ整備できることで、2019年1月予定の全国がん登録情報の提供開始後、調査研究への速やかで積極的な活用と国民への結果の還元が期待される。

## 課題1.全国がん登録を基盤とする高精度、長期持続可能な「がん診療情報」の収集の仕組み

5

- 「がん診療情報」の追加収集
  - >手始めに特定地域であれば、米国のSEER地域のように住民単位に追加情報収集可能か？
    - ユタ州がん登録等の視察で見極める
- 既存の「がん診療情報」とのリンケージ
  - >既存の何らかの「がん診療情報」とのリンケージをイメージ
    - SEER Linked Datasetsのように
- 長期持続可能な
  - >2015年以前の「地域」がん登録情報も含めてイメージ

## 課題2.「特定匿名化情報」の提供の仕組み

6

- 本人同意なく登録、登録情報の二次利用
- 信頼される匿名化
  - >SEER Data と同等のデータをSEERと同等の気軽さで提供可能にしたい
    - SEER Dataの「匿名化」レベルを確認
    - 米国の「匿名化」の考え方と一般認識を確認

## 論点1：米国のSEER地域のように住民単位に「がん診療情報」の追加情報収集可能か？

- これまでのご意見
  - 日本では現実的ではない？
    - 病院、診療所が多く、追加情報のしっかい性を担保できない
    - 都道府県には「義務」以外を行う資金がない
      - 住民単位ではないが、院内がん登録を日本版SEERとして整備するのが現実的で、需要もある。
- 〇〇であれば、できるかもしれない
  - (例)「特定の地域」で「さかのぼり」であれば、できるかもしれない
  - (例)「がんの種類限定」「項目限定」「期間限定」であれば、できるかもしれない

## 論点2：既存の何らかの住民単位の「がん診療情報」とのリンケージによる追加情報収集可能か？

- リンケージデータベースの候補
  - NDB
  - 住民単位のコホート
  - 政府統計調査
    - がん登録推進法17条に基づき国（国がん）がリンケージ後の匿名化データを整備するならば可能？

## (参考) SEER Linked Databases

- SEER-Medicare Linked Database
  - The SEER-Medicare data reflect the linkage of two large population-based sources of data that provide detailed information about elderly persons with cancer, which can be used for an array of epidemiological and health services research.
    - Visit the SEER-Medicare Linked Database Web site for more information and instructions to access these data.
- SEER-Medicare Health Outcomes Survey Linked Database
  - The SEER-MHOS database links two large population-based sources of data that provide detailed information about elderly persons with cancer. The data come from SEER's cancer registries that collect clinical, demographic and cause of death information for persons with cancer and the Medicare Health Outcomes Survey (MHOS) that provides information about the health-related quality of life (HRQOL) of Medicare Advantage Organization (MAO) enrollees.
    - Visit the SEER-MHOS Linked Database Web site for more information and instructions to access these data.
- SEER-Consumer Assessment of Healthcare Providers and Systems (SEER-CAHPS) Linked Data Resource
  - The SEER-CAHPS data set links 3 types of data: clinical (SEER), survey (CAHPS), and administrative/billing (Medicare claims and enrollment files). The data are linked through the collaborative efforts of NCI, the SEER registries, and CMS.
    - Visit the SEER-CAHPS Web site for more information and instructions to access this data.

## 論点3：SEER Data と同等のデータをSEERと同等の気軽さで「特定匿名化情報」の提供を可能にしたい

- 知見
  - 米国の定義
    - 非特定化データ- 個人情報には除かれているが、原情報に連結可能なデータ
    - 匿名化データ -原情報に連結不可能な非特定化データ
  - SEER Dataの背景
    - 研究を統括するCFR 米合衆国・連邦規則集
    - The Common Rule: Title 45, part 46
    - Title 21, parts 50, 54, 59
    - 非特定化データの利用は、ヒトに関する研究とはみなされない。
  - SEER Data の非特定データには、部分的な日付（診断日と初回治療開始日の診断の月と年）と診断時の居住地の郡が含まれる。
- 日本で「匿名化情報」として許容されるレベルは？

# 全国がん登録実装 匿名化技法一覧

手法	詳細	備考
削除	指定列を削除する	氏名、生年月日、住所列の削除など
マスキング	指定列を指定する文字列で置き換える	
レコード順の並び替え	指定列の並び替えを行う	
トップコーディング	指定列のグルーピング、リコーディングを行う	<ul style="list-style-type: none"> <li>閾値以上（以下）の数値を一般化</li> <li>小数点以下を切り捨て、範囲化</li> </ul>
曖昧化	指定列のk匿名化を行う k匿名化：データの匿名性を評価する指標	「同じような属性の人が、必ずk人以上いる状態」のこと
誤差の導入	指定列の値に誤差を導入する	
スワッピング	指定列の値を他の行と置き換える	統計的な周辺頻度分布は変えないように交換する

## 研究用提供の情報1

- がん登録推進法の定める登録対象腫瘍
- 腫瘍単位：罹患集計対象との違い

罹患集計	IARC多重がん規則 報告時の規則で集約された腫瘍 (統計対象区分=1のみ)
------	---

提供	IARC多重がん規則 登録（集約）時の規則で集約された腫瘍
----	-------------------------------

### － 主な違い

	提供	罹患集計
結腸、皮膚	亜部位単位	最初の罹患のみ
一部の局在	局在単位	報告時の規則で集約された腫瘍単位
側性	側性単位	最初の罹患のみ

## 研究用提供の情報2

基本情報	性別、 <u>診断時年齢</u> 、 <u>診断時年齢（小児用・月齢）</u> 、 <u>診断時住所コード（11桁）</u> 、 <u>診断時住所市区町村コード</u> 、 <u>診断時住所保健所コード</u> 、 <u>診断時住所医療圏コード</u> 、 <u>診断時住所都道府県コード</u>
腫瘍情報	多重がん番号、局在コード、 <u>診断名（和名）</u> 、 <u>形態コード</u> 、 <u>性状コード</u> 、 <u>分化度</u> 、 <u>組織診断名（和名）</u> 、 <u>ICD-10コード</u> 、 <u>ICD-10（和名）</u> 、 <u>統計対象区分</u> 、 <u>IARC-ICCC3</u> 、 <u>ICCC（英名）</u>
診断情報	<u>診断日</u> 、 <u>診断日精度</u> 、 <u>診断根拠</u>
進行度	<u>進展度・治療前</u> 、 <u>進展度・術後病理学的</u> 、 <u>進展度・総合</u>
治療情報	外科的治療の有無、鏡視下治療の有無、内視鏡的治療の有無、観血的治療の範囲、放射線療法の有無、化学療法の有無、内分泌療法の有無、その他の治療の有無

## 研究用提供の情報3

### • がん登録推進法の定める登録項目並びに算出項目

病院情報	<u>初診病院コード（+※保健所コード、医療圏コード、住所コード）</u> 、 <u>診断病院コード（+※）</u> 、 <u>観血的治療病院コード（+※）</u> 、 <u>放射線治療病院コード（+※）</u> 、 <u>薬物治療病院コード（+※）</u>
追跡情報	<u>生死区分</u> 、 <u>原死因（ICD-10コード）</u> 、 <u>原死因（和名）</u> 、 <u>生存期間（日）</u> 、 <u>死亡日/最終生存確認日資料源</u> 、 <u>生存率集計対象区分</u>
その他の情報	<u>発見経緯</u> 、 <u>患者移動動向</u> 、 <u>患者受療動向</u>
登録精度情報	<u>DCN区分</u> 、 <u>DCI区分</u> 、 <u>DCO区分</u>

### • 特別項目

生存率用項目	<u>生死区分</u> 、 <u>原死因（ICD-10コード）</u> 、 <u>原死因（和名）</u> 、 <u>生存期間（日）</u> 、 <u>死亡日/最終生存確認日資料源</u> 、 <u>生存率集計対象区分</u>
小児用項目	<u>診断時年齢（小児用）</u> 、 <u>IARC-ICCC3</u> 、 <u>ICCC（英名）</u>

## 定型出力

- 利用又は研究の目的に対して、必要十分な情報提供

目的	必要十分な情報
1 罹患率又は生存率の計測のみ	性別、診断時年齢、診断時住所都道府県コード、腫瘍情報
2 検診の精度管理	+診断時住所市区町村コード、発見経緯、進行度情報
3 初回治療情報を含む	+診断時住所保健所コード、診断時住所、医療圏コード、初回治療情報
4 詳細な地理的分析可能な情報を含む	+診断時住所コード、病院情報、患者異動動向、患者受療動向

## 特別な項目：病院情報

### 自動集約編集要領概略

集約初診病院コード	集約性状コード3>2>1>0 集約診断日が最古の集約初診病院コード
集約診断病院コード	1.集約診断日から4ヶ月以内の情報を対象とし、集約診断根拠昇順で優先して取得 2.1.の条件を満たすがん情報がない→空白
集約観血的治療病院コード	以下の条件に合う集約観血的治療病院コードを取得。 1.集約診断日から4ヶ月以内の外科的治療あり、鏡視下治療あり、内視鏡的治療ありデータを取得 2.集約性状コード3>2>1>0 3.外科的治療あり>鏡視下治療あり>内視鏡的治療あり 4.集約診断日が最古 5.条件を満たすがん情報ない→空白
集約放射線治療病院コード	1.集約診断日から4ヶ月以内の放射線治療ありデータを取得 2.観血的治療の2,4,5共通
集約薬物治療病院コード	1.集約診断日から4ヶ月以内の化学療法又は内分泌療法ありデータを取得 2.観血的治療の2,4,5共通



## 特別な項目：患者受療動向と患者異動動向

患者受療動向 (患者目線の受療動向の指標)	入力都道府県と診断時患者住所都道府県が 1:すべて一致 2:不一致を含む
患者異動動向 (受療の自県完結率の指標)	入力都道府県と診断時患者住所都道府県が 1:すべて同一(自県完結) 2:すべて異なる(他県依存) 3:一つでも異なる

## 病院等への提供 2019年1月～

### 届出対象情報

No	生年月日	診断施設	鏡視下治療の有無
病院等の名称	診断時住所	治療施設	内視鏡的治療の有無
診療録番号	側性	診断根拠	観血的治療の範囲
カナ氏名(氏)	局在	診断日	放射線療法の有無
カナ氏名(名)	原発部位	発見経緯	化学療法の有無
氏名(氏)	組織	進展度・治療前	内分泌療法の有無
氏名(名)	性状コード	進展度・術後病理学的	その他治療の有無
性別	分化度	外科的治療の有無	鏡視下治療の有無

### 最終生存確認日又は死亡日及び死因

死亡日	死亡を確認した場合にあっては、その死亡日
原死因	死亡者情報票に記録された死亡の原因
最終生存確認日	生存を確認した直近の日 死亡者情報票と照合後：照合年の12月31日 死亡者情報票と照合前 <sup>185</sup> 当該者に係る診断日の最も遅い日