

# がん登録のケースファインディングに有効な識別モデルの検証： 多施設共同研究

○小原 仁<sup>1)</sup>, 今村 知美<sup>2)</sup>, 岡 貴之<sup>3)</sup>, 秀平 優<sup>3)</sup>, 蔵多 嘉 陽子<sup>4)</sup>, 平岡 紀代実<sup>5)</sup>, 山口 直美<sup>6)</sup>

1) 久慈大学バイオ統計センター, 2) 横浜国立大学, 3) 高知記念病院医療管理課, 4) 高知総合病院医療情報管理課, 5) 高知医療センター企画課, 6) 神戸医療センター企画課

## 【目的】

がん登録のケースファインディングに該当する病名情報と診断や治療などの診療行為情報からがん登録症例を判別する識別モデルの開発と開発した識別モデルの判別能を複数施設の診療情報を用いて検証した。

がん登録症例を識別する統計モデルの概要



我々はがん登録に関連する病名情報と診断情報や治療情報などの診療行為情報からがん登録症例を識別する統計モデルを開発した。

公開：小原 仁, 今村 知美, 岡 貴之, 秀平 優, 蔵多 嘉 陽子, 平岡 紀代実, 山口 直美. がん登録のケースファインディングに有効な識別モデルの開発と検証. JACR 2018, 24( 3): 22.

## がん登録症例の識別モデル

先行研究で開発された識別モデルの判別能は、施設別の診療実績から得られた成果であるため、他施設でも同様の成果を得られるかは明らかでない。



A1 先行研究に用いた6施設モデルの判別能

判別結果	がん登録		
	計対象	対象	計
登録症例	5,834	35	5,869
非登録症例	0,802	0,037	0,765
登録症例	1,438	302	1,800
登録症例	0,198	0,963	0,238
計	7,272	407	7,679
	1,000	1,000	1,000

感度 0.963 (95%信頼区間: 0.940-0.979)  
特異度 0.802 (95%信頼区間: 0.793-0.811)

## 【方法】

分析対象は検証を実施した5施設において、2017年1月1日から2017年12月31日までの間に国立がん研究センターから提供されているがん登録対象ICD病名、またはがん登録情報ICD病名が付与された患者とした。病名などの診療情報とレポートデータから生成したデータセットをもとに、がん登録の有無を目的変数としたロジスティック回帰分析を施設ごとに行い、がん登録症例を判別する識別モデルを開発した。開発された各識別モデルの判別能は、感度と特異度、ケースファインディング対象症例からの除外割合をもとに評価した。



検証実施施設は多施設共同プロジェクトへの参加意向とアンケート調査をもちきり5施設とした。



## 検証実施施設の概要



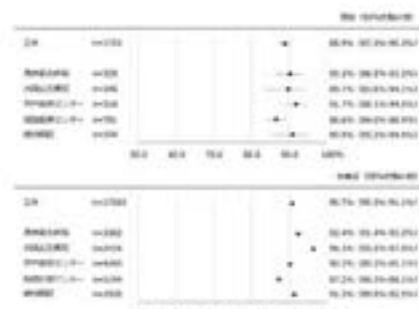
## 【結果】

がん登録のケースファインディングの対象となる件数は、施設ごとにそれぞれ7,420件、4,850件、9,313件、11,950件、4,168件であった。ロジスティック回帰分析の結果、がん確定病名や病理診断、手術などの実施情報をもつ識別モデルが開発された。各識別モデルのAUC値は、それぞれ0.964、0.970、0.946、0.917、0.956であった。評価用データに識別モデルを適用した判定結果は、感度90.2%、89.7%、91.7%、86.6%、90.9%、特異度92.4%、96.3%、90.2%、87.2%、91.3%、ケースファインディング対象症例からの除外割合85.1%、91.4%、84.7%、77.6%、85.2%であった。また、施設別の判定結果を統合した判別能 (95%信頼区間) は、感度88.9% (87.3%-90.3%)、特異度90.7% (90.3%-91.1%)、がん登録のケースファインディング対象症例からの除外割合83.5% (83.0%-84.0%)であった。

施設	感度 (%)	特異度 (%)	除外割合 (%)
施設1	90.2	92.4	85.1
施設2	89.7	96.3	91.4
施設3	91.7	90.2	84.7
施設4	86.6	87.2	77.6
施設5	90.9	91.3	85.2



## 評価用データセットを用いたがん登録症例の施設別判別能 (95%CI)



## 【考察】

複数施設の診療情報を用いて開発した識別モデルの判別能は、感度と特異度が約90%の精度でケースファインディング症例の約83%を除外した。本識別モデルは、がん登録のケースファインディングの効率化をすすめるうえで役立つ。

## 施設別の予測モデルとAUC値

施設	感度 (%)	特異度 (%)	AUC
施設1	90.2	92.4	0.964
施設2	89.7	96.3	0.970
施設3	91.7	90.2	0.946
施設4	86.6	87.2	0.917
施設5	90.9	91.3	0.956



## 本研究成果の活用事例：予測モデルを用いた未登録症例の再調査



多くの場合、自然発覚で登録されたがん登録症例がもれなく登録されているかは不明なことが多い。その際、約90%の判別精度を有する本識別モデルを活用し、未登録症例の検索を実施できる。

具体的には識別モデルで高確率に判定された症例を確認すればよい。