一般口演 | 広域保健医療・連携医療支援
一般口演9
広域保健医療・連携医療支援
2019年11月22日(金) 14:50 ～ 16:20 J会場 (国際展示場 展示ホール8・特設会場2)

[2-J-2-03] レセプト情報・特定健診等情報データベース（ NDB）を用いた死亡アウトカムの追跡

○久保 慎一郎[1]、野田 龍也[1]、明神 大也[1]、西岡 祐一[1]、中西 康裕[1]、降旗 志おり[2]、東野 恒之[2]、今村 知明[1] （1. 奈良県立医科大学公衆衛生学講座, 2. （株）三菱総合研究所 ヘルスケア・ウェルネス事業本部 ）

キーワード：NDB, Rezept, total gastrectomy, stent placement, number of patient

【目的】レセプト情報・特定健診等情報データベース(以下， NDB)とは、日本の保険診療の悉皆データである。このビックデータを活用し、死亡転帰を用いて日本全体の死亡の追跡と術後予後の評価を行った。

【方法】本研究は医療計画策定に係る評価指標作成の一環として行った。3年分の NDBレセプトを用いた。外科手術と内科治療で代表的な胃全摘術と経皮的冠動脈ステント留置術を行った患者に対して死亡割合と SMRを算出した。また、術後7カ月~18カ月の再入院回数や時間外受診者数を算出した。死亡情報は NDBの死亡転帰を用いた。

【結果】2015年度の胃全摘術を行った患者の再入院回数は(0.82回/人、0.72回/人；男性、女性。以下同)、外来の時間外受診者数は(3.92回/人、3.88回/人)であった。1年後の死亡割合は40歳以上で (15.9%、16.4%)、死亡統計の死亡率から手術1年後の死亡割合を除算した期待死亡割合は(15.9%、16.4%)、 SMRは(473、865)であった。年齢階級別に表すと60歳台まで死亡割合は横ばいであるが、70歳以上から徐々に死亡割合が高くなった。2015年度のステント留置術の再入院回数は(1.03回/人、1.05回/人)、外来の時間外受診者数は(1.46回/人、2.46回/人)であった。1年後の死亡割合は40歳以上で (4.1%、5.3%)、期待死亡割合は(4.1%、5.3%)、SMRは(129、170)であった。60歳台から徐々に死亡割合が高くなり、ピークは95歳以上であった。

【考察】胃全摘術の患者は、ステント留置術患者に比べ80歳以上の超過死亡が多く、胃全摘の80歳以上の患者の予後が低下することを意味する。ただし、リスク調整されておらず単純な比較はできない。また、患者の重症度などの背景を調整できていない。今後は両手術に限らず様々な疾患を対象に算出していく必要がある。

©一般社団法人 日本医療情報学会

レセプト情報・特定健診等情報データベース（NDB）を用いた死亡アウトカムの追跡

久保慎一郎[*1]、野田龍也[*1]、西岡祐一[*1]、明神大也[*1]、中西康裕[*1]、
降簱志おり[*2]、東野恒之[*2]、今村知明[*1]
[*1] 奈良県立医科大学 公衆衛生学講座,
[*2] (株) 三菱総合研究所 ヘルスケア・ウェルネス事業本部

Mortality tracking using the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB)

Kubo Shinichiro [*1], Noda Tatsuya [*1] , Nishioka Yuichi [*1] , Myojin Tomoya [*1] ,
Nakanishi Yasuhiro [*1] ,Furihata Shiori [*2],Higashino Tsuneyuki[*2], Imamura Tomoaki[*1]
*1 Department of Public Heath, Nara Medical University ,
*2 Healthcare and Wellness Division, Mitsubishi Research Institute, Inc.

【Background】The National Database of Health Insurance Claims and Specific Health Checkups (NDB) is a comprehensive database of the medical services under health insurance in Japan. We used this big data to track mortality and evaluate postoperative prognoses.

【Methods】This research was conducted in order to develop the evaluation criteria related to preparation of medical care plans. Using three years of NDB Data, we calculated the mortality rate and standardized mortality ratio (SMR) for patients who underwent total gastrectomy and percutaneous coronary stent placement which are popular surgical procedures in internal medicine. In addition, the number of hospital readmissions and the number of patients who were 7 to 18 months postoperative who visited a hospital out of hours were calculated. The mortality flag variable in the NDB was used for information on deaths.

【Results】The number of readmissions after total gastrectomy in 2015 was (0.82 times/person, 0.72 times/person; men, women), the number of patients who visited a hospital out of hours was (3.92 times/person, 3.88 times/person). The one-year mortality rate among patients who were more than 40 years old was 15.9%, 16.4%, expected mortality rate that mortality rate of mortality statistics divided by one-year mortality rate after surgery is 13.7%, 14.4%, SMR is 486,913. In terms of age-specific mortality rate, it remained stable until sixties then gradually increased from the age of 70. The number of readmissions after percutaneous coronary stent placement in 2015 was (1.03 times/person, 1.05 times/person), the number of patients who visited a hospital out of hours was (1.46 times/person, 2.46 times/person). The one-year mortality rate among patients who were more than 40 years old was (4.1%, 5.3%), expected mortality rate was (4.1%, 5.3%), SMR was 147,185. The age-specific mortality rate increased slowly from the sixties, and peaked at more than 95 years old.

【Conclusion】There was more excess death in patients over 80 years old who underwent total gastrectomy than patients who underwent percutaneous coronary stent placement. This means the prognosis for patients who are over 80 years old is worse. However, we cannot compare simply without adjusting the risk. We were also unable to adjust for the patient's background such as the severity of their condition. In the future, we need to calculate for a variety of diseases, not just these two surgical procedures.

Keywords: NDB, Rezept, total gastrectomy, stent placement, number of patient

1. はじめに

レセプト情報・特定健診等情報データベース(以下, NDB)とは、病院や診療所から国に送信される電子レセプトデータと特定健診等のデータを個人が特定されないように一部の情報を匿名化・削除した上で、格納・構築されているデータベースである。国民皆保険制度をとる日本における保険診療の悉皆データであり、2009 年 4 月～2018 年 3 月診療分で約 153 億 1,200 万件(2019 年 3 月末時点)[1]のレセプトデータが蓄積されるなど、世界最大級の健康関連データベースといえる。これを有効に活用することで日本全体の臨床研究(コホート研究等)や政策研究が強力に推進できると期待されている。

ただし、NDB にはさまざまな問題点や障壁が存在する。まず、政策用途を目的としたデータであるため、利用するためには申請が必要である。申請し有識者の審査を受けてデータが切り出されるには早くても半年程度の時間がかかる。そのうえ、著者らが行った先行研究では、NDB の巨大なサイズによる処理の限界と、診療報酬請求のために設定されているレセプトの構造がそのままでは研究目的での利用に適さない形式となっていることから研究者が個人で分析をするのが困難である問題点が挙げられた [2]。加えて、最も致命的なものとして NDB に付与されている 2 種類の ID が経過とともに変化しやすいという問題が挙げられる。なお、NDB には 2019 年 4 月現在で一生涯不変の個人 ID は含まれていない。代わりに、個人紐付け用の匿名変数として、「ID1」と「ID2」が用意されている。ID1 は保険者番号、被保険者証等記号・番号、生年月日、性別から個人情報保護のためにハッシュ関数と呼ばれる関数を用いて変換された英数字列であり、ID2 は氏名、生年月日、性別から同様に変換された英数字列である。ところが、

同一患者でも、就職・転職等で保険者は変化し、医療機関での表記ゆれ(例:渡辺と渡邉)や結婚・離婚等のライフイベントで氏名表記は変化するため、ID1、ID2 ともに容易に変わり得ることが分かった[3]。筆者らの先行研究により、ID1 と ID2 を組み合わせて一つの個人 ID を作成することに成功した[4)5)]。これらは NDB をコホート化する上での基盤技術となっており、さまざまな研究班でこの名寄せロジックが用いられ始めている。

患者数を集計することは可能となったが、NDB にはまだ課題が存在する。それは、アウトカム指標が乏しいことである。死亡や罹患のようにアウトカムにできる指標が乏しいため、治療の成果などが明らかにできない。

「死亡」に関してはレセプトの転帰区分の「死亡」を用いることで死亡を抽出することができる。ただし、これには①医療機関が正しく入力していること、②医療機関外の死亡が補足できることの二つが重要となる。ただし、どちらも精度が高いとは必ずしも言えない状況である。特に、医療機関の付与忘れや付与間違い等は一定発生していると推察される。

本研究は診療行為や薬剤等から死亡推定し、NDB の死亡情報の正確度向上を目的とした。また、実際の死亡データと比較してその有効性を検証した。さらに、作成した死亡フラグを用いて日本全体の死亡の追跡と術後予後の評価を行った。

2. 方法

本研究は医療計画策定に係る評価指標作成の一環として(5 年生存率など死亡関連指標のため)行った。4 年分(2013 年 4 月~)の奈良県の国保および後期高齢者レセプト(以下、KDB)と、5 年分(2013 年 4 月～)の NDB レセプトを用いた。

KDB には被保険者マスターに記載されている死亡転帰(保険者の持っている死亡情報、レセプトが発生していない患者を含む)とレセプトの死亡転帰(レセプトが発生している患者の死亡。以下医療内死亡)の2種類の死亡転帰が存在する。前者の被保険者マスターに基づいた死亡は死亡届が出された情報に基づいて作成されているため正確である。この被保険者マスター死亡をアウトカムとし、診療行為や薬剤などのレセプトの内容について機械学習を行い、レセプト死亡転帰の精度が高めることを目指した。機械学習は R 言語による決定木分析(R-part)を用いた。

決定木分析を行うに当たっては、条件設定や変数の選択が重要となる。まずは様々な情報を組み合わせて、どのような決定木を採用することが死亡転帰の精度が高まるか繰り返し決定木を作成した。説明変数には、退院日または死亡日に算定された診療行為および医薬品を使用した。診療行為は小分類で集約した。ただし基本診療料、医学管理料、在宅、精神を除外した。医薬品は薬価基準コードの先頭 4 桁で集約した。作成した決定木を基に NDB で死亡決定が行えるよう死亡推定ロジックを作成した。その仕組みを用いて NDB に適用させた後、レセプトの死亡転帰に比べてどの程度付与できたか検証した。また、人口統計の死亡数と比較しその有効性を検証した。

さらに、作成した死亡フラグを用いて日本全体の死亡の追跡と術後予後の評価を行い、臨床的指標に合致するか検証を行った。外科手術と内科治療で代表的な胃全摘術と経皮的冠動脈ステント留置術、経皮的冠動脈形成術を行った患者に対して手術後 3 年間を追跡した。各治療に対して死亡割合と SMR を算出した。SMR は 3 年間の死亡割合で計算した。術後 7 カ月~18 カ月の再入院回数や時間外受診者数を算出した。

3. 結果

3.1 決定木の分類方法について

KDB 上で医療機関が報告した死亡転帰の正確度を検証した。特異度は 99.8%と高い正確度であったが、感度が 91.4%と低かった。死亡転帰日以降にレセプトが発生しているケースは、約 2%存在し、間違って死亡とされていた。転帰区分に死亡が付与されている場合は、死亡転帰の誤りを修正するルールを作成することで、死亡の正確度を向上させる必要があることが分かった。

死亡転帰が生存となっている患者においては、機械分析を行った。レセプトにおける正例(生存)と負例(死亡)の比率の関係で、サンプリングの補正処理を行い正確度を高めた。KDB のレセプトの内、全レセプトから無作為抽出して決定木分析を行った場合、最初に分岐する項目が「食事加算」であり、入院と外来の2パターンに分かれることが分かった。また、外来において決定木にかかる診療行為を確認すると、「往診料」が最上位でその後「在宅患者訪問指導料」と「心停止」に分岐しており、往診と来院で分かれていることが分かった。よって、入院、在宅、来院の3つのモデルに分けて機械分類を実施した。これらのモデルの総称を死亡決定ロジックと呼ぶ。

3.2 KDB での入院・在宅・来院検証結果

入院、在宅、来院で決定木を作成した。代表して入院の決定木を図 1 に示す。

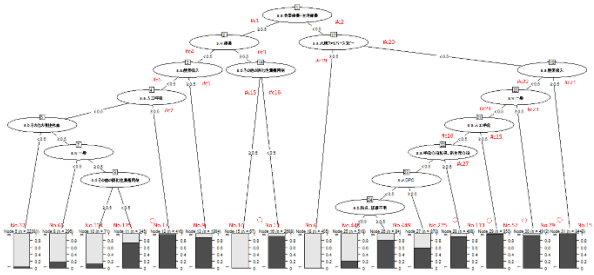


図 1　入院の決定木

決定木分析は「予測」や「判別」、「分類」を目的として使われるデータマイニング手法であり、今回は KDB のデータをランダムで2分割し、片方のデータでマスター死亡を基にした決定木を作成した。作成した決定木をもう片方のデータに当てはめた結果を抽出した。決定木の判別結果は「生存」か「死亡」かを判別して分類している。図 1 の枝葉も「死亡」と「生存」を振り分けている。また、表 1 に各モデルの検証結果を示す。入院モデルを除いてある一定の正確度が担保されたが、入院は偽陽性率が 5.2%と他に比べ高かった。

表 1　各モデル別の推定正確度

	転帰死亡あり 死亡の付与間違いを修正		転帰死亡なし 決定木ルールを適用			⑤全適用 (①~④)
	転帰死亡 そのまま	①転帰死亡 の精度向上	②入院	③在宅	④来院	
感度	91.4%	91.3%	94.7%	93.2%	94.4%	94.2%
特異度	99.8%	99.8%	94.8%	99.6%	96.9%	98.8%
偽陽性率	0.2%	0.2%	5.2%	0.4%	3.1%	0.2%
偽陰性率	8.6%	8.7%	5.3%	6.8%	5.6%	5.8%

3.3 NDB への死亡推定ルールの適用

上記3つのモデルを基として、NDB に死亡を付与した。その際、偽陽性率が高いルールは第一種過誤を避けるため除外した。死亡推定ルールを表 2 に示す。死亡適用後の患者数と人口動態の死亡数を比較した数の比を表 3 に示す。

なお死亡者の 99%は 40 歳以上であるため、NDB への適

用は 40 歳以上に限定した。

表 2　死亡決定に使用するルール

モデル	死亡推定ルール	偽陽性率
入院	最終診療日において、食事療養・生活療養なし、かつ大腸ファイバースコピー（D313）なし、かつ酸素吸入（J024）あり	0.17%
	最終診療日において、食事療養・生活療養なし、かつ大腸ファイバースコピー（D313）なし、かつ酸素吸入（J024）なし、かつ一般病棟以外、かつ人工呼吸（J045）あり	0.20%
	最終診療日において、食事療養・生活療養あり、かつ療養病棟以外、かつ酸素吸入（J024）なし、かつ人工呼吸（J045）あり	0.26%
在宅	最終診療日において、往診料（C000）なし、かつ在宅ターミナル加算なし、かつ看取り加算あり	0.00%
	最終診療日において、往診料（C000）なし、かつ在宅ターミナル加算あり	0.00%
	最終診療日において、往診料（C000）あり、かつ検体検査判断料（D026）なし、かつ処方箋料（F400）あり、かつ死亡診断加算あり	0.00%
	最終診療日において、往診料（C000）あり、かつ検体検査判断料（D026）なし、かつ処方箋料（F400）なし、かつ調剤料（F000）なし、かつ在宅患者訪問診療料（C001）なし、かつ死亡診断加算あり	0.00%
来院	最終診療日において、非開胸的心マッサージ（J046）あり	0.00%
	最終診療日において、非開胸的心マッサージ（J046）なし、かつ呼吸心拍監視（D220）なし、かつ心停止・詳細不明（I469）あり	0.08%

表 3　ルール適用後の正確度結果

年度	人口動態統計	転帰死亡精度向上		決定木ルール		全ルール適用	
		推定死亡数	(比)	推定死亡数	(比)	推定死亡数	(比)
2013年度	1,267,908	914,115	(72.1%)	92,528	(7.3%)	1,006,643	(79.4%)
2014年度	1,272,534	932,053	(73.2%)	95,542	(7.5%)	1,027,595	(80.8%)
2015年度	1,290,444	946,060	(73.3%)	113,595	(8.8%)	1,060,255	(82.2%)

3.4 NDB での推定結果の検証

人口動態調査の死亡者数と NDB 転帰区分、死亡ロジック適用分で比較したデータを代表して 2013 年度分を示す(表 4)。39 歳以下は今回死亡ロジックの対象外としたため割愛している。全体で約 8%前後の死亡の付与を行うことができた。また、医療機関外の死亡を追うことはできないため、年間 15 万人の検死事例を除外した際の値を参考値として記載した。

表 4　人口動態調査死亡と比較(2013)

年 齢	人口動態調査	転帰区分のみ		死亡ロジックでの追加		合計	
	2013	推定死亡者数	推定死亡者数/人口動態調査	推定死亡者数	推定死亡者数/人口動態調査	推定死亡者数	推定死亡者数/人口動態調査
40～44歳	10,124	5,293	52%	1,790	18%	7,083	70%
45～49歳	14,053	7,922	56%	1,646	12%	9,568	68%
50～54歳	20,581	12,103	59%	1,934	9%	14,037	68%
55～59歳	31,365	19,453	62%	2,497	8%	21,950	70%
60～64歳	62,932	40,196	64%	4,558	7%	44,754	71%
65～69歳	82,495	57,734	70%	5,953	7%	63,687	77%
70～74歳	111,419	81,540	73%	7,919	7%	89,459	80%
75～79歳	161,275	119,348	74%	11,202	7%	130,550	81%
80～84歳	224,159	168,935	75%	15,739	7%	184,674	82%
85～89歳	244,044	185,941	76%	17,629	7%	203,570	83%
90歳以上	284,653	205,602	72%	21,661	8%	227,263	80%
総数	1,247,100	904,067	72.5%	92,528	7.4%	996,595	80%
検案補正後(参考)	1,097,100		82.4%		8.4%		91%

また、国保および後期高齢者に絞って死亡患者数を比較したものを表 5 に示す(39 歳以下も含む)。

表 5 国保・後期高齢者と NDB で比較した死亡数の比率

	国保＋後期高齢者 死亡者数			NDB死亡者数			比率		
	2013	2014	2015	2013	2014	2015	2013	2014	2015
北 海 道	51,223	51,492	52,617	43,693	44,696	46,325	85%	87%	88%
青 森 県	14,300	14,162	14,570	12,045	12,177	12,579	84%	86%	86%
岩 手 県	14,682	14,805	15,049	12,079	12,354	12,679	82%	83%	84%
宮 城 県	19,836	20,239	20,449	17,134	17,551	18,283	86%	87%	89%
秋 田 県	13,027	13,155	13,059	11,695	11,929	11,916	90%	91%	91%
山 形 県	13,777	13,762	13,766	12,161	12,141	12,568	88%	88%	91%
福 島 県	21,404	21,522	21,679	18,196	18,158	18,661	85%	84%	86%
茨 城 県	26,917	27,052	27,330	22,879	23,303	24,332	85%	86%	89%
栃 木 県	18,365	18,254	18,278	15,261	15,053	15,524	83%	82%	85%
群 馬 県	19,321	19,237	19,307	17,235	17,302	17,707	89%	90%	92%
埼 玉 県	49,394	50,331	50,587	44,085	45,733	46,986	89%	91%	93%
千 葉 県	44,957	46,946	47,942	40,920	42,024	43,741	91%	90%	91%
東 京 都	92,051	93,953	93,252	74,300	77,368	79,590	81%	82%	85%
神奈川県	60,314	61,500	62,440	54,427	55,951	57,924	90%	91%	93%
新 潟 県	24,894	24,787	24,961	21,530	21,818	22,399	86%	88%	90%
富 山 県	11,592	11,573	11,625	9,356	9,525	9,761	81%	82%	84%
石 川 県	10,861	11,093	10,921	9,415	9,882	9,986	87%	89%	91%
福 井 県	8,061	8,089	8,206	6,837	6,891	7,187	85%	85%	88%
山 梨 県	8,600	8,673	8,695	7,067	7,030	7,168	82%	81%	82%
長 野 県	21,972	22,406	22,061	19,394	20,094	20,401	88%	90%	92%
岐 阜 県	19,384	19,168	19,804	16,796	16,754	17,668	87%	87%	89%
静 岡 県	34,339	34,946	35,076	29,112	29,514	30,039	85%	84%	86%
愛 知 県	52,040	52,128	52,932	48,095	49,054	51,372	92%	94%	97%
三 重 県	17,622	17,637	17,928	14,253	14,570	15,086	81%	83%	84%
滋 賀 県	10,944	10,975	10,993	9,657	9,571	9,861	88%	87%	90%
京 都 府	21,208	21,382	21,604	18,348	18,603	19,214	87%	87%	89%
大 阪 府	65,774	64,379	63,024	58,235	59,889	62,414	89%	93%	99%
兵 庫 県	47,956	48,603	49,016	40,560	41,359	42,950	85%	85%	88%
奈 良 県	12,518	12,538	12,413	10,536	10,655	10,777	84%	85%	87%
和歌山県	11,367	11,241	11,176	9,362	9,192	9,513	82%	82%	85%
鳥 取 県	6,456	6,424	6,586	5,774	5,712	6,053	89%	89%	92%
島 根 県	8,561	8,330	8,712	7,433	7,386	7,376	87%	89%	85%
岡 山 県	18,935	19,164	19,369	15,778	15,994	16,522	83%	83%	85%
広 島 県	25,956	26,752	26,615	21,486	22,433	22,924	83%	84%	86%
山 口 県	16,754	16,485	16,318	13,550	13,640	13,820	81%	83%	85%
徳 島 県	8,472	8,832	8,480	6,907	7,188	7,177	82%	81%	85%
香 川 県	10,314	10,496	10,419	9,017	9,269	9,512	87%	88%	91%
愛 媛 県	15,568	15,849	15,462	12,735	12,964	12,904	82%	82%	83%
高 知 県	8,707	8,693	8,521	7,236	7,223	7,182	83%	83%	84%
福 岡 県	41,690	42,091	42,329	35,398	36,423	37,486	85%	87%	89%
佐 賀 県	8,708	8,875	8,597	7,153	7,571	7,562	82%	85%	88%
長 崎 県	15,063	15,107	14,779	11,919	12,231	12,433	79%	81%	84%
熊 本 県	18,366	18,749	18,381	14,872	15,245	15,286	81%	81%	83%
大 分 県	12,303	12,589	12,241	10,483	10,819	10,649	85%	86%	87%
宮 崎 県	11,489	12,033	11,809	9,465	10,037	10,013	82%	83%	85%
鹿児島県	19,050	19,027	18,597	16,132	15,814	15,840	85%	83%	85%
沖 縄 県	9,006	9,125	9,046	7,583	7,861	8,052	84%	86%	89%
全国	1,096,111	1,106,663	1,109,036	939,597	959,965	987,417	86%	87%	89%

これは、KDB での死亡推定結果が奈良県固有の医療を示すものでないか検証した。死亡ロジックを作成した奈良県では 84～86%であり、他の地域も概ね 80%台で推移していた。

3.5 胃全摘術・ステント留置術・PCI 後 3 年追跡

死亡ロジックを用いて実際の治療の実態を明らかにするために基礎集計を行った。2013 年 4 月～の 2 年間で胃全摘術・経皮的冠動脈ステント留置術を受けた患者を対象に 3 年間追跡し、死亡数を算出した。各治療別の結果を表 6、表 7 に示す。なお、39 歳以下と 95 歳以上は患者数が少ないため除外した。SMR については NDB3 年分の患者数に対し 3 年間の死亡数で経年を合わせて算出した。

定義として「患者数」:手術を行った患者数、「死亡者数」:観察期間内に死亡した患者数、「死亡割合」:死亡者数／患者数、「死亡率」:生命表の死亡率(5 歳刻みの場合は中央の年齢の死亡率を採用)による 3 年間期待値、「期待死亡者数」:患者数×死亡率、「超過死亡者数」:死亡割合－死亡率とした。また、超過死亡(人口統計の死亡率から各疾患での死亡率を割ったときの死亡しやすさの指標)を「超過死亡比率 1」:((死亡者数／死亡捕捉率 0.9)／患者数)／死亡率、NDB で補足できない医療機関外で死亡した事例や検視を含む数を補正した「超過死亡比率 2」:((死亡者数／死亡捕捉率 0.9)／患者数)／(死亡率×(1-検死の割合 0.13))を算出した。

NDB の全患者の患者数、入院患者の入院日数、入院回数、再入院回数、外来患者の患者数を表 6 に示す。表 6、表 7 をもとに 3 年後の死亡割合を示したグラフを図 2～図 5 に示す。

表 6　胃全摘術後の死亡率

性別	年齢階級	患者数	死亡者数	死亡割合	死亡率	期待死亡者数	超過死亡割合	超過死亡比率1	超過死亡比率2
男性	40～44歳	156	36	23.1%	0.4%	1	22.7%	71.9	82.6
	45～49歳	269	82	30.5%	0.6%	2	29.9%	60.2	69.1
SMR=	50～54歳	513	143	27.9%	0.9%	5	27.0%	34.0	39.1
486	55～59歳	1,191	321	27.0%	1.5%	17	25.5%	20.6	23.7
	60～64歳	2,540	682	26.9%	2.3%	59	24.5%	12.8	14.8
	65～69歳	3,804	1,187	31.2%	3.8%	143	27.4%	9.2	10.6
	70～74歳	4,683	1,723	36.8%	5.7%	267	31.1%	7.2	8.2
	75～79歳	4,518	1,997	44.2%	9.1%	412	35.1%	5.4	6.2
	80～84歳	3,098	1,637	52.8%	16.1%	500	36.7%	3.6	4.2
	85～89歳	1,136	719	63.3%	27.4%	311	35.9%	2.6	3.0
	90～94歳	138	103	74.6%	42.4%	59	32.2%	2.0	2.2
	合計	22,046	8,630	39.1%		1,775			
女性	40～44歳	161	53	32.9%	0.2%	0	32.7%	171.5	197.1
	45～49歳	201	58	28.9%	0.3%	1	28.5%	98.1	112.8
SMR=	50～54歳	273	85	31.1%	0.5%	1	30.6%	69.4	79.8
913	55～59歳	443	144	32.5%	0.7%	3	31.8%	49.4	56.8
	60～64歳	775	254	32.8%	1.0%	8	31.7%	35.3	40.5
	65～69歳	1,138	395	34.7%	1.6%	18	33.1%	24.6	28.3
	70～74歳	1,294	502	38.8%	2.4%	31	36.4%	18.0	20.6
	75～79歳	1,456	604	41.5%	4.2%	61	37.3%	11.0	12.6
	80～84歳	1,100	586	53.3%	8.2%	91	45.0%	7.2	8.3
	85～89歳	539	299	55.5%	16.0%	86	39.5%	3.8	4.4
	90～94歳	112	73	65.2%	30.2%	34	35.0%	2.4	2.8
	合計	7,492	3,053	40.8%		334			

表 7　経皮的冠動脈ステント留置術後の死亡率

性別	年齢階級	患者数	死亡者数	死亡割合	死亡率	期待死亡者数	超過死亡割合	超過死亡比率1	超過死亡比率2
男性	40～44歳	4,168	135	3.2%	0.4%	15	2.9%	10.1	11.6
	45～49歳	7,713	227	2.9%	0.6%	43	2.4%	5.8	6.7
SMR=	50～54歳	11,878	383	3.2%	0.9%	108	2.3%	3.9	4.5
147	55～59歳	17,015	664	3.9%	1.5%	247	2.4%	3.0	3.4
	60～64歳	28,921	1,433	5.0%	2.3%	672	2.6%	2.4	2.7
	65～69歳	38,977	2,662	6.8%	3.8%	1,464	3.1%	2.0	2.3
	70～74歳	43,117	3,878	9.0%	5.7%	2,461	3.3%	1.8	2.0
	75～79歳	37,866	5,259	13.9%	9.1%	3,455	4.8%	1.7	1.9
	80～84歳	27,004	5,673	21.0%	16.1%	4,358	4.9%	1.4	1.7
	85～89歳	12,313	3,909	31.7%	27.4%	3,372	4.4%	1.3	1.5
	90～94歳	2,420	1,060	43.8%	42.4%	1,027	1.4%	1.1	1.3
	合計	231,392	25,283	10.9%		17,223			
女性	40～44歳	296	10未満	3%以下	0.2%	10未満	3%以下	15.8	18.2
	45～49歳	646	28	4.3%	0.3%	2	4.0%	14.7	16.9
SMR=	50～54歳	1,157	57	4.9%	0.5%	6	4.4%	11.0	12.6
185	55～59歳	2,137	88	4.1%	0.7%	16	3.4%	6.3	7.2
	60～64歳	5,079	302	5.9%	1.0%	52	4.9%	6.4	7.4
	65～69歳	9,351	520	5.6%	1.6%	147	4.0%	3.9	4.5
	70～74歳	14,107	893	6.3%	2.4%	339	3.9%	2.9	3.4
	75～79歳	16,217	1,558	9.6%	4.2%	680	5.4%	2.5	2.9
	80～84歳	15,733	2,322	14.8%	8.2%	1,295	6.5%	2.0	2.3
	85～89歳	9,853	2,387	24.2%	16.0%	1,578	8.2%	1.7	1.9
	90～94歳	3,070	1,156	37.7%	30.2%	927	7.5%	1.4	1.6
	合計※	77,646	9,311	12.0%		5,042			

表 8　NDB 全患者における入院外来受診実態

			入院				外来		
				日数	入院回数	再入院回数		時間内	時間外
	年齢階級	全患者数	患者数	日数/年/人	回数/年/人	回数/年/人	患者数	日数/年/人	日数/年/人
男性	40～44歳	3,706,571	155,389	30.4	1.28	0.28	3,699,473	7.6	1.52
	45～49歳	3,367,674	174,058	33.7	1.34	0.34	3,358,514	9.0	1.93
	50～54歳	3,119,594	201,652	36.0	1.40	0.40	3,108,648	10.7	2.43
	55～59歳	3,024,694	248,652	37.7	1.47	0.47	3,011,015	12.5	2.89
	60～64歳	3,404,247	358,809	39.2	1.54	0.54	3,384,391	14.7	3.30
	65～69歳	4,271,426	571,221	39.1	1.62	0.62	4,242,066	17.6	3.85
	70～74歳	3,182,063	551,091	38.8	1.67	0.67	3,155,293	23.0	5.06
	75～79歳	2,712,002	579,802	42.5	1.70	0.70	2,682,264	27.9	6.26
	80～84歳	2,030,451	546,090	48.3	1.71	0.71	1,993,773	31.4	7.33
	85～89歳	1,147,753	382,360	54.7	1.72	0.72	1,113,905	31.0	7.42
	90～94歳	411,761	161,765	58.9	1.70	0.70	393,455	28.0	6.68
	合計	30,378,236	3,930,889	42.4	1.61	0.61	30,142,797	16.9	3.79
女性	40～44歳	4,230,503	208,826	21.6	1.27	0.27	4,226,745	9.2	1.71
	45～49歳	3,812,300	172,005	27.7	1.34	0.34	3,807,638	10.2	1.96
	50～54歳	3,488,545	175,628	33.3	1.42	0.42	3,482,172	12.0	2.46
	55～59歳	3,329,932	194,570	36.3	1.48	0.48	3,322,045	13.4	2.83
	60～64歳	3,736,264	267,263	38.9	1.50	0.50	3,724,353	15.3	3.29
	65～69歳	4,739,928	423,157	40.1	1.52	0.52	4,720,975	18.6	4.11
	70～74歳	3,730,688	442,053	41.4	1.54	0.54	3,710,357	24.9	5.76
	75～79歳	3,382,062	528,167	47.8	1.58	0.58	3,355,813	29.8	7.19
	80～84歳	2,926,610	612,595	56.3	1.60	0.60	2,883,354	32.2	8.04
	85～89歳	2,104,258	572,174	65.0	1.62	0.62	2,048,149	30.3	7.41
	90～94歳	1,124,825	378,987	71.5	1.61	0.61	1,073,711	26.7	5.92
	合計	36,605,915	3,975,425	48.3	1.54	0.54	36,355,312	18.8	4.25

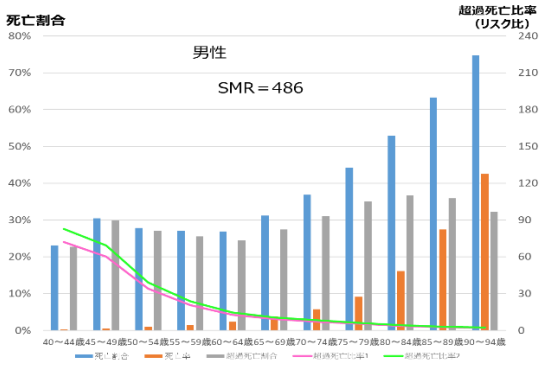


図 2　胃全摘術後の 3 年後死亡割合(男性)

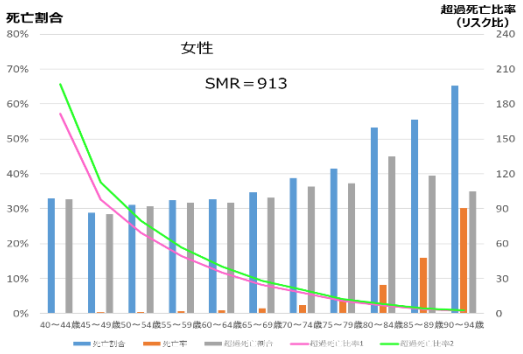


図 3　胃全摘術後の 3 年後死亡割合(女性)

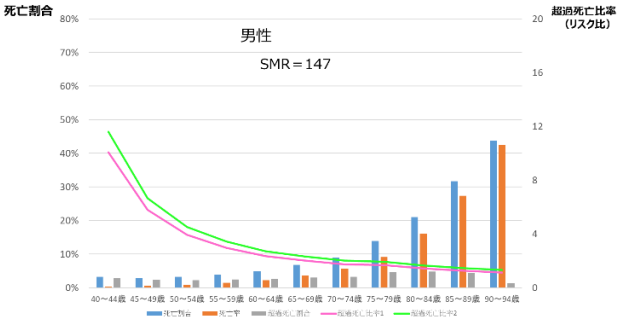


図 4　ステント留置術後の 3 年後死亡割合(男性)

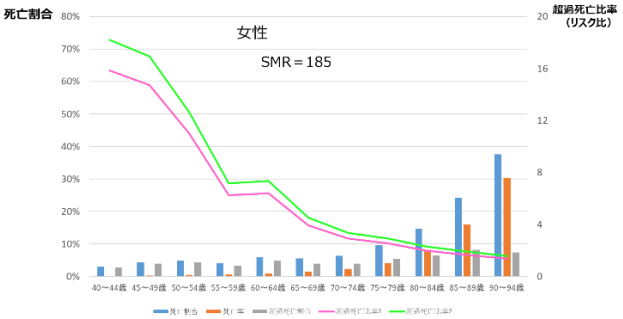


図 5　ステント留置術後の 3 年後死亡割合(女性)

4. 考察

4.1 NDB におけるアウトカム指標の重要性

NDB には患者の悉皆性が確保されている一方で、前述の通りアウトカム指標が乏しいという問題がある。それはレセプトという医療費の請求のためのデータを格納しているためであるが、逆にいえば医療費請求に必要な情報しか網羅されておらず、電子カルテに記載されているような「確定診断」「検査などの補助情報」「ADL などの状態の変化」「各種観察内容」は追うことができない。唯一補足可能な転帰死亡も精度が高い状況とは言えない。表 3 を参照すると、転帰死亡では 7 割程度しか死亡を補足できないのである。このような状況(図 6)のなかで如何に死亡情報を使えるものとするか、については NDB が使えるようになるかの重要な事項といってもいいであろう。NDB の課題についてはさまざま存在したが、図 6 のとおり 6 つの課題の内半数を解決している。そして、今 4 つ目を改善するところである。死亡アウトカムの補足が可能となれば、患者の死亡率や(期別)生存率を明らかにすることができる。どの薬剤が生存期間を延長することが可能かの追跡も可能になるほか、薬剤の未知の知見も明らかになるであろう。

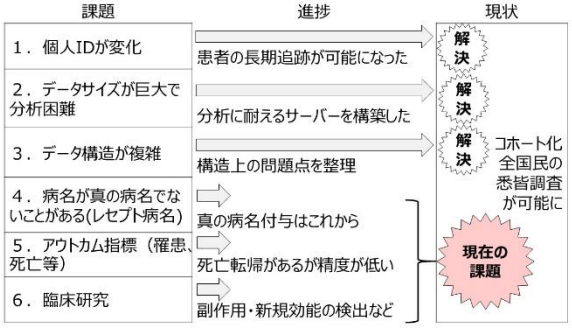


図 6　NDB の課題に対する進捗

4.2 死亡数の検証

死亡の補足が可能な患者を増やしても、間違った付与を行っては死亡ロジックの死亡転帰を使用することはできない。今回の死亡ロジックでは、偽陽性を極限まで低下させる(間違った死亡を付与しない)ことで死亡ロジックの信頼性を確保させることとした。

人口動態統計の死亡者数を患者死亡の全数として比較すると、転帰区分のみの死亡では 73.3%の正確度であったものを 82.2%まで高めることができた。人口動態統計には医療機関外で死亡したものも含まれているため、医療機関で補足可能な死亡の大部分を補足可能となったと考えられる。例えば、2013 年においては、警察庁の発表している検案患数を医療機関外死亡と仮定すると 91%まで補足が可能となったことが示される(表 4)。

しかし、NDB での死亡ロジック適用後の死亡数が正しいか正確に検証することは容易ではない。推定死亡数を向上させたが、前述のとおり医療外死亡が一定存在し、医療機関に受診しないまま死亡している人が一定数存在するためである。警察の検視事例の中に医療機関に搬送されて死亡が確定した場合も含むと想定されるため、確実な検証はできない。ただし、検案の大部分が医療機関外で起こっていることであろうことは違和感がない。今後は精緻化が必要であろう。

ロジックの適正化については決定木が奈良県特有の医療を示すものではないことを証明するために表 7 で都道府県によって大きく乖離がないか確認した。NDB の都道府県は医療機関所在地のため、都心部は若干数が多く出ているところがあるものの、概ね 80%台で推移しているところを見ると全体の特徴をある程度捕捉できているのではないかといえる。ただし、これらをより詳細にみるには他の都道府県のレセプトを用いて比較することが必要であるといえる。

4.3 決定木の検証と臨床での信頼性

決定木の条件が臨床に合うものでなければ懐疑的と言わざるを得ない。これまでに多くのトライアンドエラーを繰り返しながら決定木を作成してきた。たとえ AUC や感度、特異度が高くても臨床的に説明できないモデルは不採用とした。入院であれば「最終日に食事をしていない人が内視鏡検査を受けたかどうか」という条件が存在する。これは絶食中でも内視鏡検査の場合は検査のみ受けて帰宅する(＝予後もいい)ということの表れであるといえる。在宅の場合は、在宅訪問診療料や、死亡診断加算、看取り加算などが説明変数として抽出された。来院に関しては、心臓マッサージや、心停止、心不全、ショックといった病名が説明変数として抽出された。これらの診療行為や薬剤、病名は臨床として違和感なく、今回の決定木の説明変数としては可用性が高いといえる。

これらをもとに偽陽性を極限まで低下させる(間違った死亡を付与しない)ことでロジックの信頼性を確保させることとした。特に、全データで KDB での死亡ロジックとマスター死亡を比較した際は感度が高くても NDB の死亡ロジックと死亡転帰では感度が下がっており、必ずしも感度が高いからと言ってすべての条件をロジックに組み込むことは有用ではない。このようなことから、今回の死亡ロジックの全体像を示したものを図 7 に示す。条件によって分岐するフロチャートで死亡としたケースを図式化した。例えば、40 歳以上で、転記死亡が付与されている患者で、転記死亡日以降にレセプトが発生していない場合を死亡としている。これは、付与間違いを訂正するためのロジックである。

なお、これら死亡ロジックの条件は必ずしもこれだけではないということに留意する必要がある。分析できるデータ数が増加すれば、決定木の枝葉を細分化させることにより、より有用な死亡ロジックに組み替えることも可能である。これは今後 KDB のデータ数を増やし、他の地域の KDB で同様の検証を行うことによって実現可能である。

また、死亡ロジックでは死亡したあとにレセプトが発生する事例を確認し補正しているが、今回確認した事例以外にも存在する可能性がある。特に歯科においては一定の数が生じていると考えられており、今後の精緻化が必要である。

説明変数の選択は、種々の診療行為、医薬品、病名等から判別する必要がある。レセプトの最終日の診療行為発生に着目した理由は、最終日の処置が死亡の傾向をつかむうえで重要と判断したためである。検証段階で、集計対象に含める診療行為を期間別(最終日、3 日以内、5 日以内)で検証したが、期間が長くなれば長くなるほど余計なノイズ（日々の治療）が含まれた。最終日に限定したところが正確度を高めた重要な点である。

医学管理料および基本診療料を除外した理由については、すべての患者で継続的に発生しており説明変数としての役割が果たせなかったためである。このような変数はほかにも一定存在すると考えられる。

上記のことより、決定木を作成する上でたたき台となる説明変数を用意することができたと考えるが、今後の精緻化は必要であるといえよう。

4.4 死亡ロジックを用いた実症例での応用

死亡ロジックを用いて、基礎集計を行った。外科手術と内科治療で代表的な胃全摘術と経皮的冠動脈ステント留置術、経皮的冠動脈形成術を行った患者に対して手術後 3 年間を追跡した。

胃全摘術の患者は、ステント留置術患者に比べ超過死亡が多く、胃全摘術の患者の予後が低下することを意味する。この超過死亡は死亡する可能性の高い疾患がどこまで治療後の効果が図れるかを検証するための指標としており、図 2 ～図 5 を見ると胃全摘術患者はどの年齢も 30%ほどの超過死亡があり、死亡しやすい状況を示している。一方で、ステント留置術患者は 10%以下で推移しており、胃全摘術患者に比べると予後がいいという結果になった。がん患者は基本死亡しやすく、ステント留置術患者や経費的冠動脈形成術患者は、治療によって完治していることが数値から考えられる。また、年齢が上昇すればするほど超過死亡が大きく上昇するわけではないことも実態を示していると考えられる。

なお、SMR は女性のほうが高い傾向を示しているが、これは男性に比べて女性のほうが長寿命であることや、喫煙率が低いなどの状況が影響していると考えられる。SMR は男女差を見るよりは、全体と比較したときの差や年度の比較で用いることが有用といえよう。

各疾患別の診療実日数・回数の比較に関して、すべての疾患は年齢が上昇すると入院日数が上昇する結果となった。外来の受診日数においても同様の傾向となった。再入院の回数に関しては期間における再入院を拾っており、必ずしもその疾患で再入院となったことを示しているわけではないことに注意を要する。疾病の発生のしやすさについては期間有病率等を用いることが有用であろう。

集計に関しては医療を示す指標としては違和感がない結果といえる。ただし、他のデータとの比較を行うことができない。今後は、実際に集計したデータが他のデータで示された傾向を示すかも含めて検証が必要であろう。

4.5 本研究の限界

文中でも示したが、今回の死亡ロジックの付与によってできるだけ間違いがないように感度を重視せず特異度ができるだけ高くなるように設計している。その関係で、一部死亡とできてない事例が存在する。これは疫学視点で間違った推論を生まないように考慮した結果である。

また、今回は対象を 40 歳以上の患者に限定した。これは先述のとおり、39 歳以下の患者では死亡の絶対数が少なく、周産期と小児が合わさって出現することから、正しいツリーを導けないと判断したためである。今回は分析の限界と言わざるを得ないが、今後小児や産褥期に特化した検証が必要であるといえる。

また、死亡ロジックを用いた実症例での応用として胃全摘術とステント留置術患者の超過死亡に関して検証を行ったが、リスク調整されておらず単純な比較はできない。また、患者の重症度などの背景を調整できていない点に留意する必要がある。

5. 結語

KDB を用いて死亡患者の動向を機械学習し決定木分析を行うことで NDB の死亡フラグを策定した。死亡情報の正確度を高めることが可能となり NDB を用いたコホート研究が前進する。

謝辞

本研究は平成 30 年度厚生労働科学研究費補助金（地域医療基盤開発推進研究事業）「地域の実情に応じた医療提供体制の構築を推進するための政策研究」、平成 28,29 年度国立研究開発法人日本医療研究開発機構（AMED）地域横断的 ICT 活用医療推進研究事業「レセプト等の大規模電子診療情報を活用した薬剤疫学研究を含む医療パフォーマンス評価に関する研究」、平成 28,29 年度国立研究開発法人日本医療研究開発機構（AMED）臨床研究等ICT基盤構築・人工知能実装研究事業「新たなエビデンス創出のための次世代 NDB データ研究基盤構築に関する研究」、平成 30 年度文部科学研究費助成事業（科学研究費補助金）基盤研究（A）18H04126「データ科学・疫学・臨床医学の融合による日本の保険診療情報（NDB）の全解析」、平成 30 年度文部科学研究費助成事業（科学研究費補助金）若手研究　18K17390「全診療報酬情報を用いた糖尿病診療の時空間疫学分析」、平成31年度文部科学研究費助成事業（科学研究費補助金）若手研究　19K19464「NDB（レセプトデータ）を用いた難病患者の疫学調査の手法開発と治療実態の把握　」の一環として実施したものである。

参考文献

1) 第三者提供の現状について（報告）（成果物集計・承諾案件）, 2019. [https://www.mhlw.go.jp/content/12401000/000520125.pdf (cited 2019-August-27)].
2) 久保 慎一郎, 野田 龍也, 明神 大也, 加藤 源太, 今村 知明. NDB(ナショナルデータベース)の課題および留意点と今後の展望. 医療情報学連合大会論文集, 36(1) : 272-275, 2016.
3) 久保 慎一郎, 野田 龍也, 明神 大也,東野 恒之, 松居 宏樹, 加藤 源太, 今村 知明.レセプト情報・特定健診等情報データベース (NDB)の臨床研究における名寄せの必要性と留意点. 日本健康開発雑誌 2017;38:11-18.
4) 野田龍也、久保慎一郎、明神大也、西岡祐一、東野恒之、松居宏樹、他. レセプト情報・特定健診等情報データベース（NDB）における患者突合（名寄せ）手法の改良と検証. 厚生の指標. 64(12), 7-13, 2017-10.
5) Kubo S, Noda T, Myojin T, et al. National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB): outline and patient-matching technique. bioRxiv. [https://www.biorxiv.org/content/early/2018/04/02/280008.full.pdf. (cited 2019-August-27)].

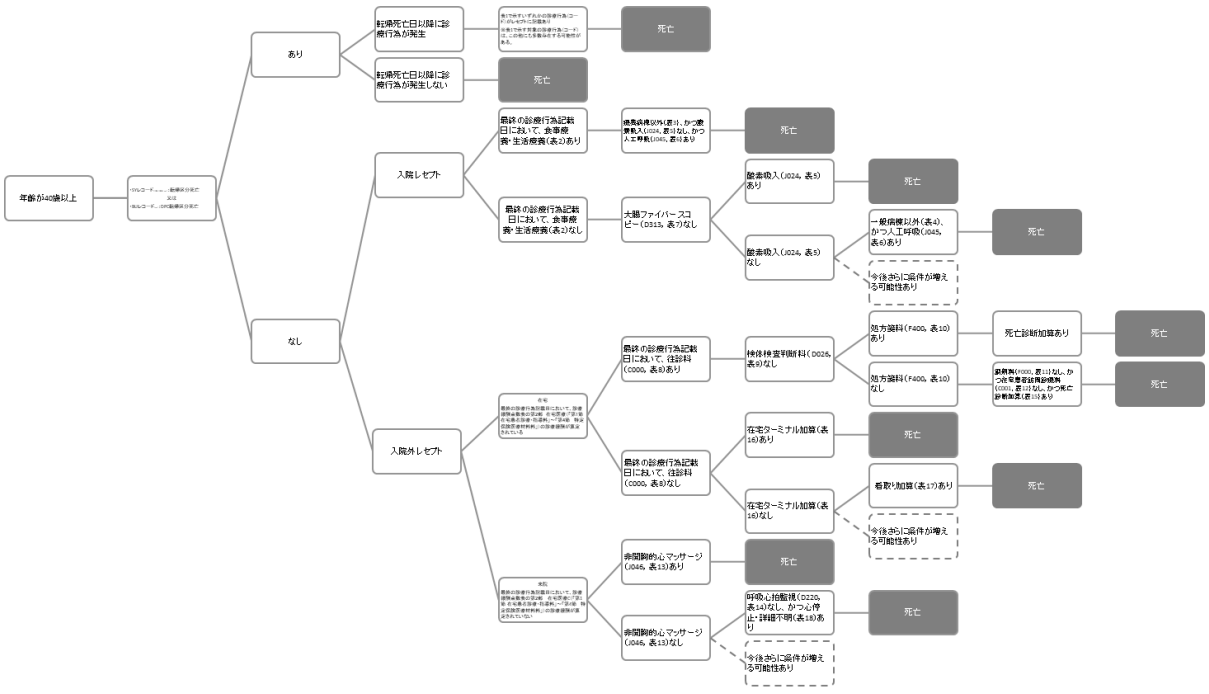


図 7　死亡ロジックの全体像