

# 医療・健診・介護データの統合化による 生活習慣病予防システムの開発

(研究助成金 60万円)

札幌医科大学 医学部 公衆衛生学講座 教授 大西 浩文

[1996年 札幌医科大学医学部卒]

共同協力者 札幌医科大学 医学部 公衆衛生学講座 助教 小山 雅之  
札幌医科大学医療人育成センター物理学教室兼 附属総合情報センター講師  
高塚 伸太郎  
札幌医科大学 医学部 公衆衛生学講座 研究生 山口 徳蔵

## 〔研究応募書〕

### 研究目的

#### 【研究の背景】

近年、超高齢社会へと進展した我が国では、疾病の発症予防のための環境整備としての「健康づくり」の重要性と必要性が高まってきている。「未病」の段階から個人の変化を的確に把握し、疾病の発症を未然に防ぐことが求められるが、現行の保健予防事業の現場で、即時に利活用でき、個人の健康に向けた行動変容の動機づけに繋がるようなデータを提供するための医療・保健・福祉データを高度に統合し分析を可能としたシステムは存在しない。国民健康保険連合会が提供しているKDBでは、介護の経年変化の分析、地域の特性に即した課題分析を行う上での自由度に制約があり、地域社会の変化に即応した施策を十分に展開できていないのが現状である。

#### 【研究成果の達成目標】

本研究では、市町村が保有する住民の医療・介護のレセプトデータと健診データに係るデータベース化を図り、市町村においてデータの分析結果を保健予防事業の展開に即時的に活用し得るプロトタイプシステムの開発を通じて、住民の健康意識の向上に資する分析結果の提供を可能とし、将来の社会実装を見据えたプラットフォームの整備につなげることを目標とする。これまで我々は健康の維持増進施策を効果的に推進できる汎用性の高いシステム開発と実践を目指して、医療サービスの需給面から見た医療格差、人口や疾病構造の変化を見極めた医療サービス提供体制の改善の必要性に関して検討を行ってきた。これには、レセプトデータを対象に

個人情報保護のためハッシュ化をした上で名寄せを行い、時系列的に、かつ、国民健康保険（国保）と後期高齢者医療制度（後期）のデータのシームレスな連結化を図った。解析結果は、データ提供を受けた市町村へのフィードバックを行い施策展開の参考に供した。本研究においては、医療のレセプトデータに加え、介護と健診データを加えた、名寄せ、串指し、紐付けを市町村独自の統一番号で実現を図ることとしており、新たに開発する統合分析システムによって、市町村における保健事業の企画・運営に資する分析結果の提供が可能となるか検証することを目標とする。さらに、将来的な目標として、地域特性に応じたシステムのカスタマイズ化と、Personal Health Record (PHR) としての健康アプリケーションとの連携も見据えた開発を検討していく。

## 研究実施計画の概要

- (1) 市町村が保有する被保険者の医療・健診・介護の個人データを匿名加工し、市町村が活用している住民基本台帳をベースにした総合行政システム番号を利用して個人で複数発生する情報を個人単位にまとめる「名寄せ」や健診・医療・介護の複数のデータベース間での同一人物の「紐づけ」を行う。このシステム番号の扱いは市町村により異なるので本研究のフィールドは予備調査の結果からE町をモデル地域として選定する。対象者はモデル地域内に住む住民で国保及び後期の被保険者とする。また、データは、E町で格納されている医療・介護レセプトデータ及び特定健診データをE町庁舎内で匿名加工が行われ、札幌医科大学への提供を受ける。
- (2) エビデンスに基づく保健事業の展開に資する分析を可能とする統合データ分析システムの考案  
保健事業の立案・展開に資する分析結果を提供する新しいデータ統合分析システムの構築に向けて、個人毎に医療・健診レセプトデータや介護データを統合化（名寄せ・紐づけ）し、健診を受けた者や保健予防事業に参加した者との医療費や、介護保険への依存度との相関関係を明確化でき、生活習慣病の発生予防・重症化予防の保健事業の施策の企画・立案、評価に有用なシステムを企画・開発する。
- (3) 提供を受けた健診・医療・介護データのデータベース化  
医療・介護レセプトデータや検診データから、未受診・未健診との医療費や介護度等を個別的に把握するためのデータベースを整備する。
- (4) PCへの実装から社会実装へ  
健診・医療・介護のデータベースを統合することによって、市町村の国保・保健担当スタッフが簡便に利用でき、ニーズに合わせた自由度の高いデータ分析が可能となるシステムを構築し、実際に対象市町村のスタッフに利用してもらうことで社会実装を目指す。
- (5) 倫理面の配慮  
札幌医科大学倫理委員会の承認を得て行われている（承認番号2-1-59）。関係法規の遵守と特に個人情報保護に留意し、保険者の同意のもと、システム開発に取り組んでいる。今回使用するデータは、提供元市町村において個人を特定できる情報は削除され、また名寄せに必要な個人識別符号はハッシュ化を行った上で札幌医科大学に提供されており、札幌医科大学においては個人を識別する一切の情報は管理をしていない。また、既存資料を活用する研究であり、研究についての説明と問い合わせ窓口を本学ホームページに公開している。

# I 緒言

全ての人々が可能な限り最高の健康水準を維持し健康寿命を延伸できることが望ましい。しかし、近年における新型コロナウイルス感染症の拡大のような脅威に見舞われることによって、個人や地域社会にとっては日常的な健康維持の困難性も浮き彫りになってきている。このような、新型コロナウイルス感染症の流行下における非感染性疾患（NCDs）に対処する事業の重要性とともに、不健康な食事や運動不足、喫煙、過度の飲酒などの原因が共通する生活習慣の改善による疾病予防の緊要性も指摘されてきている。さらには高齢化の進展と共に今後予想される医療資源の効率化も重要な課題とされている。

医療レセプトデータ（診療報酬請求書）は、患者毎、診療月毎、入院・外来・調剤別に分けて作成されていることから、個人の受療歴の追跡ができ、また、介護レセプトデータには介護給付費明細書、介護予防・日常生活支援総合事業費明細書及び給付管理票（福祉用具購入費、住宅改修費など市区町村が直接支払う費用（償還払いに係るものは除く。))が、すべて集計されている。これら医療、介護のレセプトデータは、診査機関である北海道国民健康保険連合会（以下「国保連」）の審査を経て、その都度、市町村の端末機に送信されている。

更に、特定健診・特定保健指導のデータについては、高齢者の医療の確保に関する法律（「高確法」）による実施基準により、測定、検査、審査、指導等が計画的に行われており、健康の保持増進に向けた施策が展開されている。

これらに係るパーソナルヘルスデータは、NDB（ナショナルデータベース）や、KDB（国保データベース）としても整備されているものの、個人の医療・介護・健診データを結合して未病段階からの健康維持増進に係る時系列の詳細な分析結果を得るには必ずしも適しておらず、保険者が地域の特性に応じて、いつでも自由にデータを扱い、即時的に施策展開に利活用するには、時間的、制度的な制約を伴っているのが実状である。

他方、市町村が独自に、パーソナルヘルスデータを統合化し、医療・保健・福祉の相互関連に着目したデータを利活用するための、データの入手や管理の自由度は高いので労力や時間的負担は軽く、法令の範囲内でパーソナルデータの統合化による利活用できる環境にあるもののこれらのデータを統合して分析して各種保健事業に活用している事例は少ない。これら三種の統合化したデータの分析によって得られる現状と課題から、地域住民の総合的な健康推進につながり、その成果がモデルとなり他の類似地域への展開も期待される。

本研究では、生活習慣病やメタボリックシンドロームに着目し、市町村の協力を得て医療や介護のレセプトデータ及び、特定健診並びに保健指導データに係るパーソナルデータを名寄せ・紐付けして分析が可能となるシステムを開発することにより市町村における健康課題の特定や保健予防事業の効果的な展開への活用の可能性について検討を行った<sup>1)</sup>。医療・介護・予防（＝特定健診）野路を総合的に収集・集積、分析するための情報基盤がほぼ確立しているが、必ずしも活用されてこなかったとの見解がある<sup>2)</sup>。

## II 研究方法

### 1. 対象地域と対象者及び対象期間

#### (1) 対象地域と対象者

北海道内のE町の協力を得て実施し、対象者は、国民健康保険（以下「国保」）の被保険者、及び北海道後期高齢者医療広域連合（以下「後期」）の被保険者並びに介護保険の認定者である。E町の住民基本台帳に登録されているおよそ8千人のうち、国保と後期の被保険者の割合は約4割であり、近年後期の被保険者数が増加傾向にある（表1）。

特定健診の対象者は、ほぼ人口に比例し、減少傾向にあるものの受診率は、30%台の横ばい状況にあり、受診率の分析評価の重要性が指摘されている。介護保険の認定者の年平均人数では、約550人前後で推移している（表1）。

表1 被保険者に係る主な指標の推移

(単位 人)

年度	人口	医療 被保険者		特定健診			介護保険			
		国保被保険者	後期被保険者	特定健診対象者	特定健診受診者数	受診率 (%)	要介護認定者(年平均)	要支援認定者(年平均)	合計(年平均)	
平成27	8,466	2,134	1,281	1,451	485	33	439	136	575	
28	8,235	2,010	1,257	1,374	439	32	428	145	573	
29	8,235	1,846	1,270	1,317	435	33	413	116	529	
30	8,084	1,778	1,486	1,267	375	30	423	100	523	
31	7,732	1,708	1,341	1,212	439	36	430	106	536	
出典⇒	道市町村課	資料提供—道国保医療課					資料提供—江差町高齢あんしん課			

#### (2) 総医療費と糖尿病受診歴者の医療費

平成27年度から5か年間のE町における国保と後期の総医療費は約22億円～21億円規模で、一人当たり年間医療費では61万円～66万円台で推移している（図1）。また、受療者のうち糖尿病（レセプトデータ上の「主傷病名」に記載ある者の受診者一人当たりの年間医療費は、約35万円～23万円）で減少傾向にある（図2）。

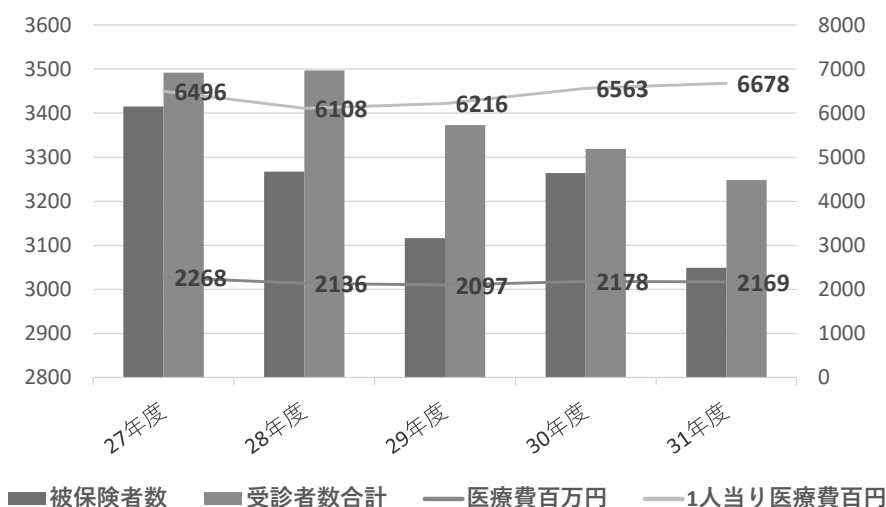


図1 国保と後期の総医療費

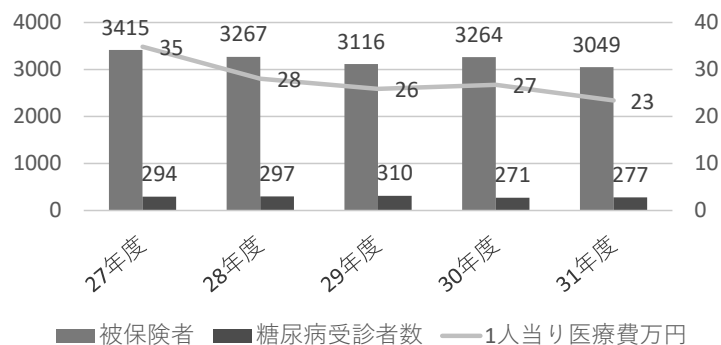


図2 65歳以上の糖尿病受診者の推移

(3) 対象期間

対象期間は、2015年度（平成27年度）～2019年度（平成31年度）の5か年間とした。

2. 対象データの種類の種類と各種データの統合

(1) データの種類

対象期間内の被保険者に係るデータは、以下の3種類を対象とし統合化した。

- ① 医療データ
- ② 介護データ
- ③ 健診データ

①には、「国保と後期の医療レセプトデータ」に分類、②には、介護給付サービスに係る詳細なデータが掲載されている「介護レセプトデータ」と要支援・要介護認定に係る情報等が盛り込まれている「介護宛名コード（E町独自の呼称）データ」に分類され、③には、特定健診、後期の健康診査、及び各々の保健指導に係るデータに分類される。

また、①と②のレセプトデータについては、国保連の審査が終了後、毎月、保険者の端末機に送信されるデータについて匿名化処理した後に分析可能なシステム構築を図った。

(2) 各種データの統合と分析方法

3種類のデータは、②のデータの「介護宛名コード」と称されている内容に住民基本台帳上の住民に附番されているハッシュ化された個人番号を活用して各データ間を紐づけし、名寄せを行った。この住民基本台帳に依拠した個人番号の性質は、E町が独自に付番されているものであり、E町の行政目的に用いられているものであることから、転出入等の移動の際に、これらのデータは、移行はされず、同一市町村内の住所を有する期間に限られる。例えば、国保から後期へ移行した場合でも同一市町村内では追跡することは可能であるものの、他の市町村間との転出入の場合は困難であり、一部制約を伴うものである。しかし、この結合方法においては、仮に同性同名でも、固有の個人番号に依拠していることから、名寄せ、紐づけの障害にならず、その精度と信頼性は確かなものとして、扱うことを可能としている。

(3) 分析方法

開発した統合データ分析システムを活用して、3種類のデータを統合することによって検討が可能になる分析を実施し、保健事業への活用の可能性を検討するため、以下の分析を行った。



- (イ) 初年度を2015年度に設定し、医療レセプトデータから判定される初年度における65歳以上の糖尿病受診歴のある者で、介護データにより判定される2019年度における介護サービスの受給の有無に与える要因の分析を行った。
- (ロ) 2015年度の糖尿病受診歴者の医療費について5分位の階層に分類した。
- (ハ) 5年後の介護認定の有無を従属変数としてロジスティック回帰分析を行った（IBM SPSSを使用）。

### Ⅲ 研究結果

#### 1. 医療・健診・介護のパーソナルヘルスデータの積集合人数と項目数

平成27～令和元年度の各年度のレセプト等データの個人番号による名寄せ後の人数の5か年分の合計人数を算出した。これらのデータの積集合人数のうち医療、介護、健診データいずれのデータも有している実人数は674人であった（図3）。

医療レセプトについては、医療機関が異なる毎に、レセプトは発生し、また同一医療機関内でも診療科が異なる毎に発生するが、これらにつき5か年を通して名寄せ後の人数としてカウントしており、他のデータも同様に捕捉している。この積集合人数については保健・医療・介護の3種類のデータが存在し、それらを活用した分析を行うことができると判断される。健診データやレセプトデータを用いて、被保険者の疾病構造や健康問題等の課題分析を行った。地域が有する保健医療資源の実態把握の重要性が指摘されている<sup>3)</sup>。

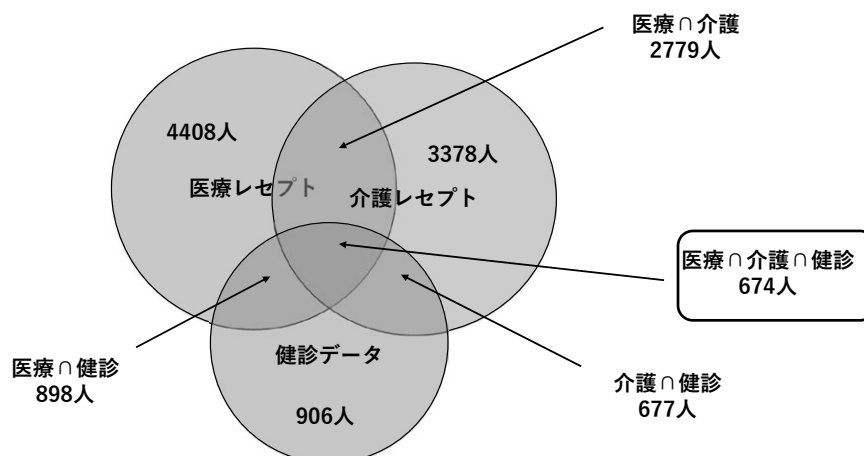


図3 医療・健診・介護のパーソナルヘルスデータの積集合人数

#### 2. 平成27（2015）年度の介護認定の有無と5年後の変化

平成27年度65歳以上の者で、糖尿病受診歴がある者のうち、5年後も転出がなく在住し、かつ介護データの存在する156人について、介護保険の要支援又は要介護の認定の有無と5年後の認定状況

の変化の実態について、27年度認定無しから5年後も認定無しというような変化を検討すると、改善したものは1人であり、悪化は17人、変化なしは121人であった（表2）。

表2 平成27（2015）年度の介護認定の有無と5年後の変化

27年度65歳以上総数	27年度要介護無	27年度要支援有	27年度要介護有
156	128	10	18

27年度65歳以上総数	31年度要介護無	31年度要支援有	31年度要介護有
156	113	9	34

27年度65歳以上 糖尿病患者	31年度要介護無	31年度要支援有	31年度要介護有	計
27年度要介護無	110	6	12	128
27年度要支援有	2	3	5	10
27年度要介護有	1	0	17	18
計	113	9	34	156

27年度65歳以上 糖尿病患者	31年度要介護無 + 31年度要支援有	31年度要介護有	計
27年度要介護無 + 27年度要支援有	121	17	138
27年度要介護有	1	17	18
計	122	34	156

### 3. 平成31年度の介護認定発生の有無での初年度（平成27年度）の対象特性の比較

平成31年度の介護認定の発生の有無の2群間で初年度（平成27年度）の対称特性の比較を行ったところ平成31年度の介護認定ありの群において平均年齢と要支援認定有の頻度、医療費の高額者の頻度が有意に高い結果であった。男性の割合と健診受診ありの頻度に関しては2群間では有意な差は認められなかった（表3）。

表3 平成31年度の介護認定悪化の有無での初年度（平成27年度）の対象特性の比較

H27年度データ		H31年度介護認定なし (n=121)	H31年度介護認定あり (n=17)	p値	使用した検定方法
年齢	mean±SD	74.2±6.3	79.4±8.0	0.003	unpaired t-test
男性	n (%)	51 (42.1)	7 (41.2)	1.000	Fisher's exact test
要支援認定あり	n (%)	5 (4.1)	5 (29.4)	0.003	Fisher's exact test
医療レセ年合計点数5分位	n (%)				
	≤ 4153	24 (19.8)	2 (11.8)		
	4154-7156	27 (22.3)	2 (11.8)		
	7157-11843	33 (27.3)	2 (11.8)	0.033	Fisher's exact test*
	11844-22637	24 (19.8)	4 (23.5)		
	22638 ≤	13 (10.7)	7 (41.2)		
健診受診あり	n (%)	16 (13.2)	0 (0.0)	0.219	Fisher's exact test

\*2×2以外はSPSSではできないのでSTATAで計算

### 4. 平成31年の介護認定有無を従属変数としたロジスティック回帰分析

H27年度糖尿病の受療歴者（65歳以上）を対象とし、5年後（平成31年度）の介護認定の悪化の有無を従属変数としてロジスティック回帰分析を行った。

(1) 単変量解析

年齢が高いほど、また要支援認定者であること、医療費が高いカテゴリに該当する場合、5年後に介護認定状況が悪化しやすいことが明らかになった (p<0.03)。

(2) 多変量解析

変数増加法による変数選択を行った分析では、H27年度における要支援認定のオッズ比が12.35であり、医療レセプト年間の合計点数5分位の最も高いカテゴリの最も低いカテゴリに対するオッズ比は7.92であった。(表4)。

表4 H31年の介護認定有無を従属変数としたロジスティック回帰分析

変数	単変量ロジスティック回帰			多変量ロジスティック回帰		
	オッズ比	95%信頼区間	p値	オッズ比	95%信頼区間	p値
年齢(歳)	1.12	1.04-1.21	0.004			
男性	0.94	0.34-2.69	0.961			
要支援認定あり	9.67	2.45-38.22	0.001	12.35	2.62-58.20	0.001
医療レセ年合計点数5分位(%)						
≤ 4153	1.00	-	-	1.00	-	-
4154-7156	0.89	0.12-6.81	0.91	1.13	0.13-9.69	0.915
7157-11843	0.73	0.10-5.53	0.76	0.67	0.08-5.75	0.711
11844-22637	2.00	0.33-11.97	0.45	2.28	0.33-15.57	0.401
22638 ≤	6.46	1.17-35.74	0.03	7.92	1.24-50.63	0.029
健診受診あり(%)*	0.28	0.00-1.80	0.219			

\*exact logistic regression

変数増加法による変数選択

5. 生活習慣改善への動機付けに向けたビジュアル化機能システムの実装

市町村の政策現場では、保健師等の職員が、自治体が保有するパーソナルデータを保険者として法令の範囲内において自由に分析できる。これらのデータは、医療・介護サービスを提供した日の1～2か月後には審査機関の審査を経て、市町村に備えられている端末機に直接送信される。このように、新規データが毎月蓄積されることから、保健師の保健指導等において、被保険者とコミュニケーションを図る上で必要と思われる資料のビジュアル化として①検索機能(図4-1)、②個人サマリー(図4-2)、③全体サマリー(図4-3)などの可視化できるアプリケーションの開発をおこなった。今後の社会実装に向けて、E町の現場スタッフより当該アプリケーションを現場で活用することによる今後の期待についての意見聴取を行ったところ、以下のような意見が得られた。

- i) 介護サービスの予測に当たって、過去の介護サービスの受給に至った者の医療と介護サービスの内容を中長期の期間設定を行って分析することにより、医療と介護のサービス内容が変化する過程やその関連性、傾向から将来の介護サービス受給に係る予測の可能性を高め、介護計画や介護保険財政の運営に反映可能であること。
- ii) 医療レセプトデータからは、国保と後期の連結と時系列的な把握を行い、高額医療費、年齢階層別分析結果のビジュアル化によって、自治体での政策立案のための根拠となる資料としての活用が期待できること。
- iii) 健診データについては、受診勧奨のための基礎資料として、例えば、HbA1c、尿酸窒素の値から



保健指導の必要度を勘案した対象者の絞り込みを行い、保健指導後のモニタリングを続ける上で、実装化したビジュアル化システムの利用価値への期待は大きいこと。

- iv) 検索機能では、期間設定や傷病名を特定した上で絞り込み、健診データや、介護データとの紐づけを行い、検索結果の一覧表示化を図っている（図4-1）。個人サマリーについては、個人の検診結果や医療レセプトから抽出した医療費や傷病名を時系列的に並列表示の方法を採用しているため、必要とする年次の個人毎の健診結果やレセプトデータからの受診歴について検索機能を利用して、抽出した者への保健指導する際に、対象者の健康状態をグラフ表記するなど（図4-2）、保健指導対象者の理解が得やすく、生活習慣の改善への動機付けに必要な資料整備の容易になること（図4-1）。しかし、保険者（市町村）の多様なニーズに応える分析を行うためには改修の余地があること。

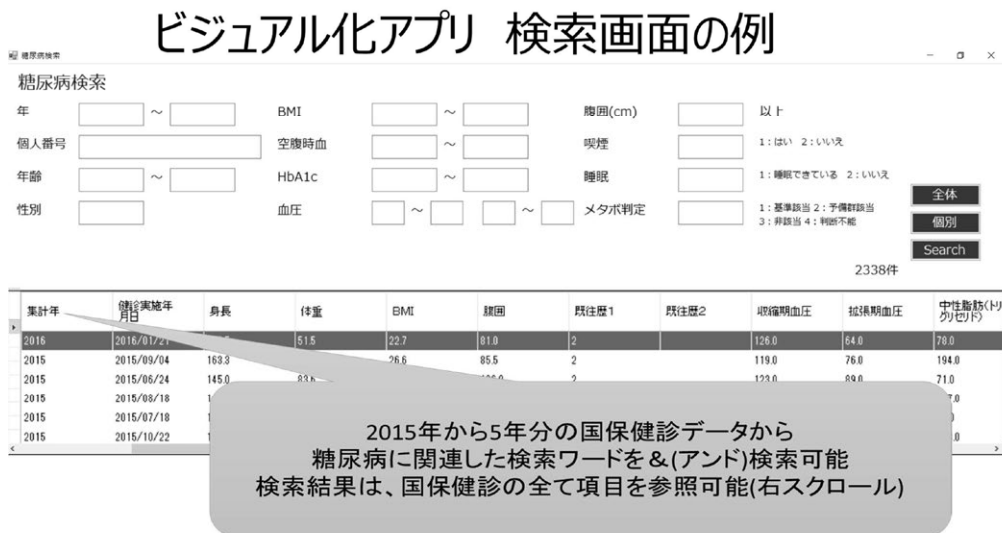


図 4-1 作成したデータベースの検索機能画面

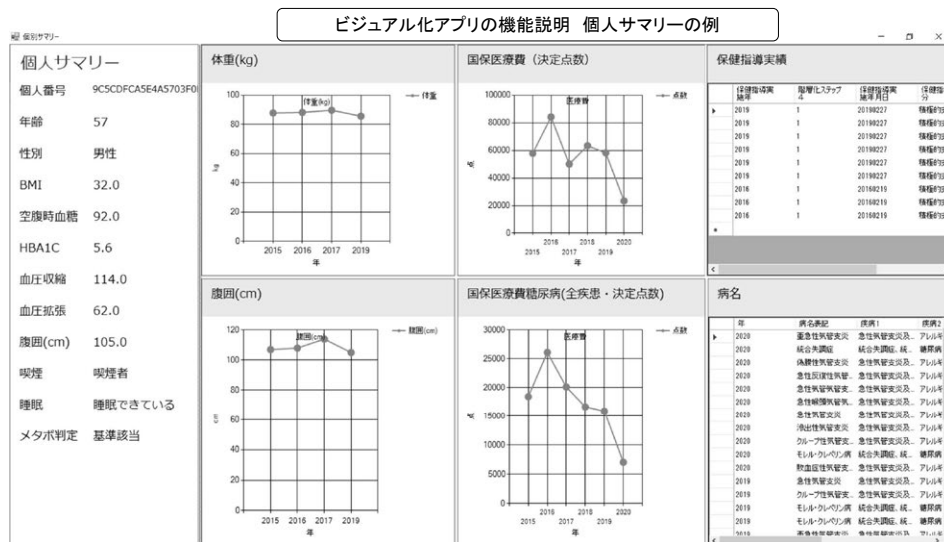


図 4-2 作成したデータベースの個人サマリー画面

v) 全体サマリーについては、ポピュレーション的発想に基づく資料整備の仕組みが取り入れられており、国保会計や介護計画の運営に関連した委員会へ提出する指標等の資料の整備や、報告書類の取りまとめに資するデータの抽出機能の機動性を図ることができる（図4-3）。

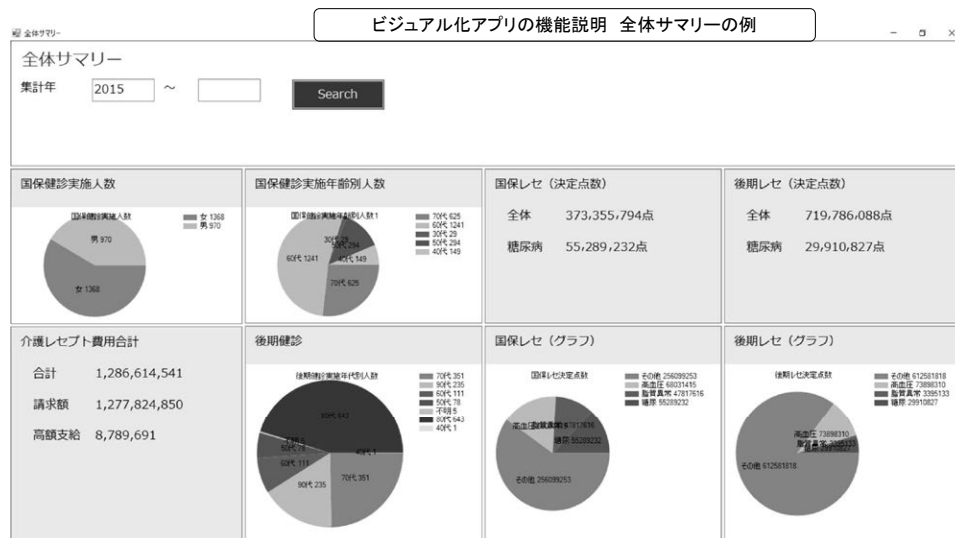


図 4-3 作成したデータベースの全体サマリー画面

## IV 考 察

保健・医療・介護の3種類のデータの統合化は、E町独自の総合行政システム番号と位置づけられている個人番号を用いることにより、数か年にわたり精度と信頼性の高い結合の実現を図ることができた。

このことから、健康から未病段階を経て医療や介護サービスの変化に至る過程の明確化に寄与し得るものとする。今回の研究では、目的変数に介護認定の有無を設定したロジスティック回帰分析の結果に着目したが、健診の未受診者への受診勧奨の結果や保健指導に係るアウトカムの評価を通じて未病段階における課題を特定し、より効果的な保健事業展開につながる可能性が考えられた。

また、個人に着目した名寄せ・紐づけによる、ハイリスクアプローチに視点を置いた検討の重要性に鑑みて、今後は、中長期的な時間軸を考慮した検討や、生活習慣の改善へのアプローチの事業の成果を検討し、効率的な施策展開の可能性をより明確化していくことが必要と考えられる。特に、保険者は、逐次、蓄積されるビッグデータを利活用できる環境にあることから、パーソナルヘルスデータの統合化は、一次予防に必要な情報や疾病のハイリスク群の把握に活かすことができる。保険者は予防事業に必要とする情報を保有するデータの分析評価の資料を以って施策展開に資することができると思われる。

政策現場における保健師等の取り組みにおいて、ビジュアル化した、資料を用いた対面における保健指導においても、これらの資料は、生活習慣の改善に向けた未病段階からの行動変容を促す動機付けへの有用性が期待され、効果的な施策展開の可能性が示唆された。

## V 結 語

本研究において、北海道のE町を対象に、保健・医療・介護の3種類のデータを統合するシステムを構築し、要介護認定に与える医療要因の影響等の分析や、保健指導に資する個人単位での健診データの時系列表示、全体のサマリーをグラフ表示することで町全体の健康度を評価することが可能であった。これらの機能は、市町村の効果的な保健事業の推進に役立つ可能性が示唆された。中長期的な評価への適用、より現場のニーズに即した視覚化の方法など、本システムのさらなる改善が今後の課題である。

## VI 謝 辞

本研究への助成を賜りました（公財）総合健康推進財団及び本事業の推進に当たり、ご理解ご支援いただいた江差町並びに関係者の皆様に深く感謝申し上げます。

### 【参考文献】

- 1) 大西浩文, 辰巳治之, 高塚伸太朗, 樋室伸顕, 小山雅之, 山口徳蔵. 道路交通網を基盤とした医療資源の適正配置に係る研究. 開発こうほう2020; 12: 20-25.
- 2) 松田晋哉, 藤野善久. 医療保険・介護保険レセプトと特定健診データの連結分析システムの開発. フィナンシャル・レビュー2015; 123(3): 84-91.
- 3) 日本医師会, 日本糖尿病対策推進会議, 厚生労働省. 糖尿病性腎症重症化予防改定プログラム. <https://www.mhlw.go.jp/content/12401000/program.pdf> (2021年11月30日アクセス)