

研究課題 84

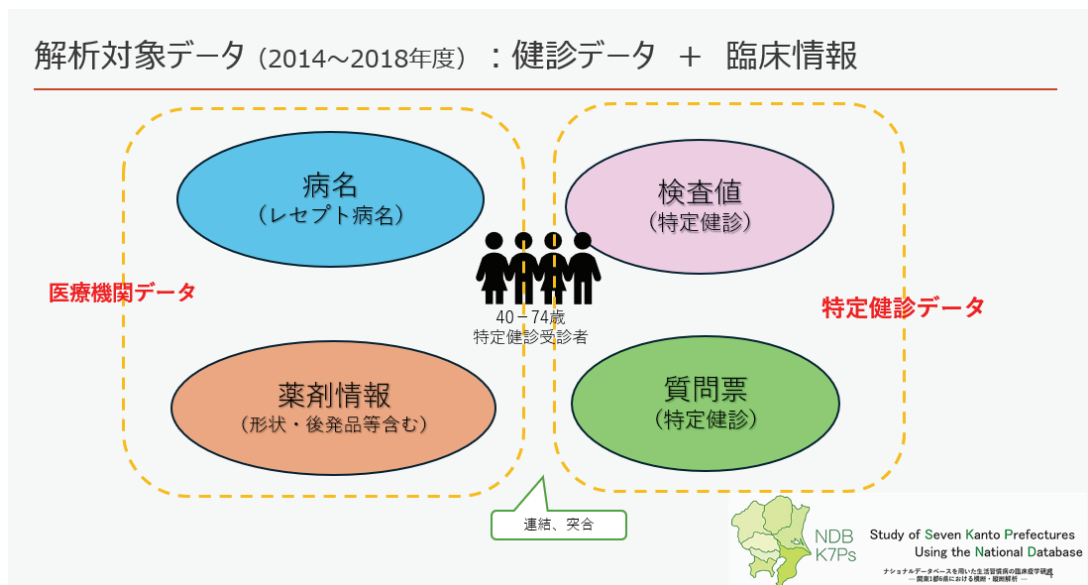
ナショナルデータベースを用いた生活習慣病の臨床疫学研究 —関東1都6県における横断・縦断解析—

食科学部栄養学科 中島 啓、関根 愛莉

1. はじめに

我々は、「ナショナルデータベース（NDB）を用いた生活習慣病の臨床疫学研究—関東1都6県における横断・縦断解析—」という研究課題に対して、厚生労働省の承認を受け、2008-2018年の特定健診、レセプト病名（高血圧、糖尿病、脂質異常症などの生活習慣病、癌など）、投与薬剤（生活習慣病の薬）の各データシートなどを2022年7月に厚生労働省から提供して頂いた。

解析するデータは、下記のような形で連結し、一つのデータシートを各年度作成した。



今までは、データクレンジングとデータシートの連結作業に多大な時間がかかっていたが、2024年6月に、高性能のPCを総合研究所の助成により導入することができ、安定したデータシート間の連結、クレンジング作業が可能となった。またデータ解析中も、メモリー不足による作業中断・フリーズ等の問題はほぼ解消した。2025年6月に、さらにもう1台の高性能PCを総合研究所の助成により導入でき、解析は円滑になり、解析結果の保存・再解析も安定した。とくに、病名（医科レセプト）データと薬剤情報データの容量が非常に大きい（それぞれ数TB）ため、このことによって、多くの疾患と薬剤の影響等を同時に検討することが可能になった。このことにより昨年度から下記を含めた複数の研究が進行中であり、一部を学会発表し、現在は論文作成中、または投稿中である。

2. 逆流性食道炎と2型糖尿病、肥満、併存疾患、および治療薬との関係

— 関東1都6県 NDB の活用 —

中島 啓

(第122回日本内科学会総会・講演会 2025年4月19日発表)

【目的】

逆流性食道炎（GERD）は、肥満、2型糖尿病、併存疾患、薬剤等によって誘発・悪化することがある。しかし、逆流性食道炎と体重、血糖コントロール、併存疾患、薬剤（副作用等）等との関係は十分に解明されていない。そこで我々は関東1都6県のNDB（リアルワールドデータ）を用いて検討した。

逆流性食道炎と体重、血糖コントロール、併存疾患、薬剤（副作用等）との関係との関係を2018年度の特健診受診者6,875,981人において調査した。

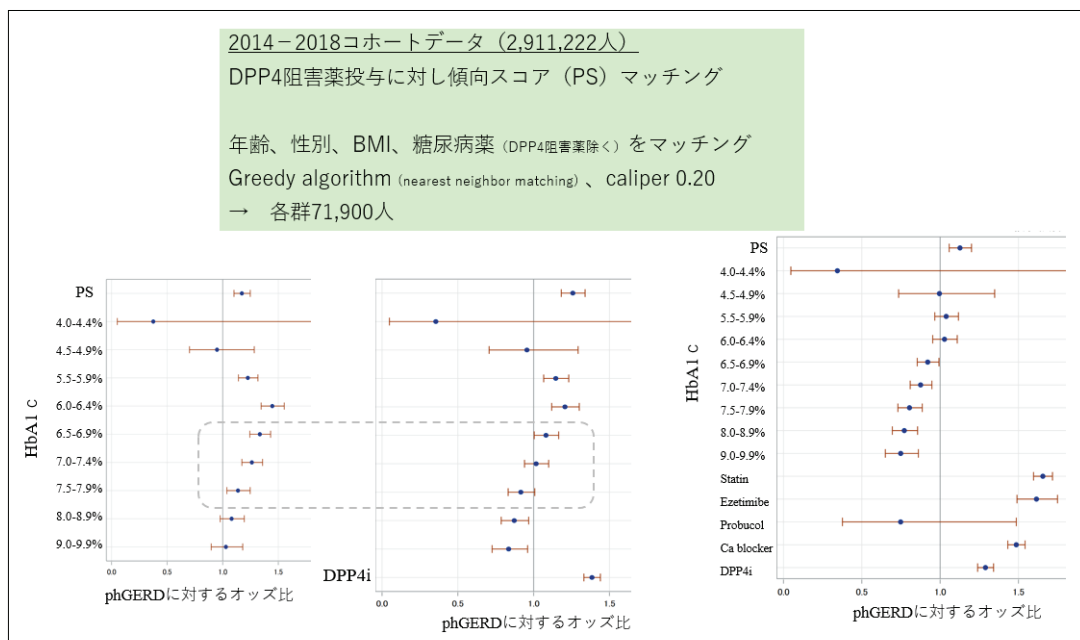
【方法】

逆流性食道炎（phGERD）を従属変数として、共変数（BMI、HbA1c、併存疾患、投与薬剤等）との関連を2018年度特定健診受診者6,875,981人において横断解析した。

SAS-EGを用いてロジスティック回帰分析し、関連の強さはWald-chi square値を参考にした。

【結果】

逆流性食道炎患者では、年齢・性を考慮しても、高血圧、糖尿病、脂質異常症、鼻炎等の有病率が逆流性食道炎のない人に比べ2～3倍であった。体重とU字型関係にあった。また高HbA1cにおけるGERDは、糖尿病の合併症、あるいは薬物療法（DPP4阻害薬）の副作用の可能性が考えられた。傾向スコアマッチングにより、ベースラインの背景を合わせた2014-2018のコホート研究においても同様なことが確認された。現在、再解析を行い論文作成中である。



3. 日本人中高年の併存疾患と処方薬の特徴：

チャールソン併存疾患指数による層別化

関根 愛莉

【目的】

チャールソン併存疾患指数（CCI）とは、死亡に寄与する 19 の併存疾患を評価し点数化した指標であり、短期的および長期的な死亡リスクと相関があるとされ、多くの疫学研究において、患者の生存率の評価や、様々な疾患の予後予測のために用いられている。しかしながら、CCI の基準に含まれない生活習慣病や感染症、眼科および耳鼻科疾患、皮膚科疾患および筋骨格疾患、精神疾患などの疾患や、処方薬との関係については不明な点が多い。NDB を用い、40～74 歳の日本人における診断病名の有病率と処方薬の使用割合について、CCI 分類と年齢層別に網羅的に明らかにすることを目的とした。

【方法】

2018 年度に特定健診を受診した関東 1 都 6 県在住の 40～74 歳を対象とし、我々が提供を受けた NDB データに含まれる医科レセプトデータ（15916 種の病名）、薬剤レセプトデータ（16886 種の処方薬）を用い、CCI を計算した（算出方法は、下記の公表研究論文 1 にて公表済み）。対象者を CCI = 0（併存疾患なし群）と CCI ≥ 4（高リスク群）に分類し、さらに、対象者の中で若年者層の 40-44 歳および高齢者層の 70-74 歳に着目し、併存疾患および使用薬を比較検討した。

【結果】

CCI 分類および年齢層別の併存疾患は、CCI の分類と年代に関わらず、アレルギー性鼻炎などのアレルギー疾患や、急性気管支炎などの一般的な感染症が上位であった。

70-74 歳では CCI の分類に関わらず、高血圧症の有病率が最も高く、高コレステロール血症など、その他の生活習慣病の有病率も高かった。しかし、CCI ≥ 4 群では、40-44 歳においても高血圧、糖尿病、高脂血症の有病率が高く（25%以上）、併存疾患の多い高リスク患者では、若年だとしても生活習慣病を併存していることが分かった。

また、特徴的な疾患の有病率について、ボンフェローニ補正による多重比較を行った結果（下表）、CCI = 0 の 70-74 歳よりも、CCI ≥ 4 の 40-44 歳の方が、下記の疾患において有病率が高いことが明らかになった。

処方薬の使用割合については、どの CCI の分類と年代においても、解熱鎮痛剤、去たん剤など一般的な感染症に関連する薬剤、アレルギー用薬の使用割合が高く、併存疾患の結果と一致していた。また、生活習慣病関連薬では、70-74 歳において、脂質異常症に対する薬剤（ロスバスタチンナトリウム錠）の使用割合が高かった。しかし、CCI ≥ 4 群の 40-44 歳では、高脂血症の有病率が高かったにも関わらず、ロスバスタチンナトリウム錠の使用割合が 6.1%と低かった。

【考察】

併存疾患の多い若年者では、高齢者と同様に、高血圧、脂質異常症の有病率が高いが、薬物療法を受けている者の割合が低かったことから、生活習慣病に対する薬物療法の有無は年齢に依存すると示唆された。よって、併存疾患の多い若年者に対しては、薬物療法が必要なレベルまで重症化し

CCI 分類、年齢層における疾患有病率の比較

病名	CCI=0		CCI ≥4	
	Age 40-44 n = 1,161,138	Age 70-74 n = 458,479	Age 40-44 n = 10,417	Age 70-74 n = 104,115
2型糖尿病	2,389 (0.21)	7,182 (1.57)	2,027 (19.5)	25,688 (24.7)
高血圧症	33,087 (2.85)	150,073 (32.7)	4,156 (39.9)	75,156 (72.2)
高脂血症	17,031 (1.47)	59,789 (13.0)	2,631 (25.3)	42,307 (40.6)
鉄欠乏性貧血	29,392 (2.53)	5,678 (1.24)	2,113 (20.3)	14,385 (13.8)
甲状腺機能低下症	4,937 (0.43)	3,992 (0.87)	769 (7.38)	5,779 (5.55)
難治性逆流性食道炎	6,193 (0.53)	15,361 (3.35)	1,547 (14.9)	23,890 (23.0)
胃炎	66,416 (5.72)	40,727 (8.88)	1,767 (17.0)	19,308 (18.5)
便秘	28,894 (2.49)	38,816 (8.47)	2,718 (26.1)	38,508 (37.0)
下痢症	28,295 (2.44)	7,809 (1.70)	1,151 (11.1)	6,928 (6.65)
うつ病	22,103 (1.90)	7,318 (1.60)	870 (8.35)	5,660 (5.44)
不眠症	36,507 (3.14)	50,311 (11.1)	2,416 (23.2)	31,490 (30.3)
不安神経症	11,596 (1.00)	11,815 (2.58)	684 (6.57)	8,418 (8.09)

All differences between age and CCI categories were significant (all p values < 0.007; according to χ^2 test with Bonferroni correction for multiple comparisons)

ないよう、食事・運動療法によるアプローチが重要であると考えられる。また、今回の解析により、CCIのような従来の死亡予測指標では見落とされてきた併存疾患（アレルギー性鼻炎や急性気管支炎、高血圧、鉄欠乏性貧血、精神疾患など）や併用薬が、死亡率予測や疾患予後に関与する可能性が示唆された。

NDB を用いて併存疾患や使用薬を網羅的に明らかにする本手法は、死亡率や疾患予後に寄与する見落とされてきた多様な因子を明らかにし、疾病のメカニズムや有効な治療方法の解明に寄与すると考えられる。本研究結果について、現在、論文投稿中である。

4. 今後の研究の展望

今後の我々の研究の目的は、大規模保健医療データを用いて、生活習慣病をはじめとする疾患の関連因子を抽出し、その影響度により優先順位を定量的に評価し序列化することである。高血圧、糖尿病、脂質異常症などの生活習慣病は、遺伝的素因、食生活、喫煙・飲酒、身体活動、検査値、肥満・痩せ、ストレス、睡眠、社会経済的要因など、多種多様な因子が複雑に関与している。しかし従来の研究では、「どの因子から介入すべきか」という優先順位が明確でない場合が多く、多因子介入の必要性は指摘されつつも、実践的な指針は十分に示されてこなかった。

我々の研究の特徴

大規模保健医療データ

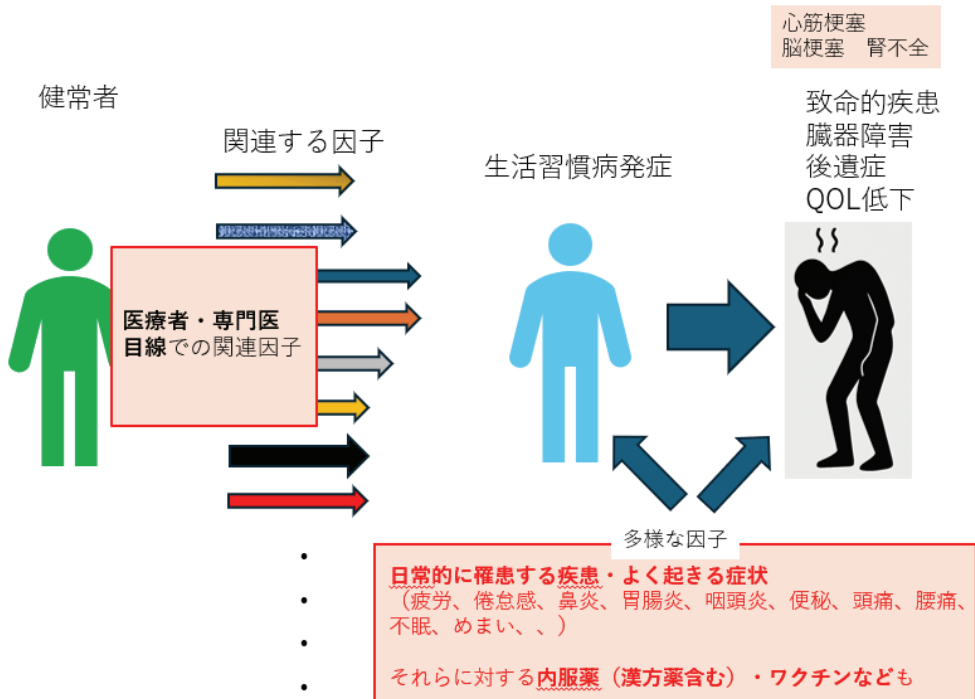
・多くの関連リスク因子の序列化

従来の統計解析（ロジスティック回帰、一般化線形など）によるWald chi-square値、標準化 β 、集団寄与危険度

機械学習による特徴量重要度、SHAP値

などによってリスク因子（関連因子）を序列化（優先順位の示唆）

本研究では、ロジスティック回帰や一般化線形モデルによる Wald カイ二乗値、標準化 β 、集団寄与危険度といった従来型統計指標に加え、機械学習による特徴量重要度や SHAP 値を用いて、関連因子の序列化を試みている。実例として、正常体重者 61 万人超を対象に、10 年後の体重 10% 以上増加に寄与する因子を Wald カイ二乗値で序列化した研究や、心不全患者の再入院リスク因子、糖尿病予測因子などの解析結果を示し、統計解析と機械学習の結果が高い相関を示すことを確認している。一方で、従来の専門家視点に基づく解析では、関連する検査値中心となりやすく、日常的な症状や服薬、ワクチン接種といった多様な因子が十分に考慮されていない課題も浮き彫りになった。



本研究では、鼻炎や便秘、胃薬・解熱鎮痛薬の使用など、一見関連が薄そうな因子も含めて評価し、疾患発症や予後への寄与を可視化する点に特徴がある。最終的には DAG（有向非巡回グラフ）を用いた因果推論により、交絡因子や媒介因子を考慮した影響評価を行い、媒介割合や除去割合を算出する計画である。今後は、2015～2024 年度の特健診、レセプト、薬剤、コロナ感染、ワクチン、死亡情報を統合解析し、研究成果の社会実装を目指す。

5. 公表研究論文（総合研究所助成を受けた 2024 年 4 月以降）

1. Airi Sekine, Kei Nakajima. Development of an algorithm for the Charlson Comorbidity Index using the National Database Study in the Japan Kanto 7 Prefectures. medRxiv 2025.04.10. 25325568.
2. Nakajima K, Sekine A. Correlations Between Wald Chi-Squared Values and Standardized Beta-Coefficients in Regression Analysis and Feature Importance Calculated Using Machine Learning. AI in Clinical Medicine, 2025. 1, e10.
3. Nakajima K, Sekine A. Association between primary hypercholesterolemia, including familial hypercholesterolemia, and abdominal aortic aneurysm in general population: Results of NDB-K7Ps-Study-5. Current Aging Science in press.

日本語要約：原発性高コレステロール血症（PH）や家族性高コレステロール血症（FH）と大動脈瘤との関連は不明であった。本研究は日本のナショナルデータベースを用い、40～74 歳約 900 万人を解析した。その結果、PH および FH はいずれも、交絡因子調整後も腹部大動脈瘤（AAA）と有意に関連していた。一般集団においても、PH と FH が特に AAA のリスク因子であることが示された。