

令和 5 年 6 月 8 日現在

機関番号：33920

研究種目：若手研究

研究期間：2019～2022

課題番号：19K17970

研究課題名（和文）隠れマルコフモデルによる糖尿病発症モデル構築と糖尿病発症予防への展開

研究課題名（英文）Hidden Markov Models for onset of type 2 diabetes

研究代表者

成定 明彦（Narisada, Akihiko）

愛知医科大学・医学部・講師

研究者番号：20800965

交付決定額（研究期間全体）：（直接経費） 2,300,000円

研究成果の概要（和文）：本研究では、隠れマルコフモデルを労働者の健康診断データに適用して、糖尿病発症モデルを構築した。モデルは、未病（健康）の状態、ハイリスク状態、糖尿病の状態を決め、状態から状態へ、毎年行き来（遷移）があるものとした。モデル構築の結果、「ハイリスク状態」の推定分布（ $104.6 \pm 7.1$  mg/dl）から、2型糖尿病のハイリスクを110 mg/dLとしているWHOのカットオフポイントは高すぎる可能性が示唆された。さらに、「ハイリスク状態」から「正常状態」への移行確率の推定値が0.01%と非常に低いことから、「ハイリスク状態」に至った者は「正常状態」に戻ることはほとんどない可能性が示唆された。

研究成果の学術的意義や社会的意義

健康診断縦断データに機械学習モデルを当てはめた結果、これまでの「ハイリスク」基準が高すぎることで、「ハイリスク状態」に至った者は「正常状態」に戻ることはほとんどない可能性が示唆された。したがって2型糖尿病予防のためには、特定保健指導の文脈とは別に「ハイリスク状態」として特定された個人に対する毎年のモニタリングと継続的な介入が必要であること、「ハイリスク状態」に至らないための「正常状態」の者たちへのポピュレーションアプローチが重要になることが推察された。このように、機械学習による2型糖尿病発症モデルは、発症の実態と予防戦略についての理解を向上させる可能性がある。

研究成果の概要（英文）：In this study, a hidden Markov model was applied to workers' health examination data to construct a model of diabetes development. The model determined the unwell (healthy) state, the high-risk state and the diabetes state, and assumed that there is an annual back and forth (transition) from state to state. The results of the model construction suggested that the estimated distribution of the 'high-risk state' ( $104.6 \pm 7.1$  mg/dl) suggests that the WHO cut-off point of 110 mg/dL for the high-risk of type 2 diabetes may be too high. Furthermore, the estimated probability of transition from 'high risk state' to 'normal state' was very low at 0.01%, suggesting that those who have reached 'high risk state' are unlikely to return to 'normal state'.

研究分野：産業保健

キーワード：糖尿病予防 健康診断 機械学習 前糖尿病

## 様式 C-19、F-19-1、Z-19（共通）

### 1. 研究開始当初の背景

(1) 糖尿病の発症予防は重要であり、なお大きな課題である。糖尿病の「ハイリスク者」に対する生活習慣改善介入が発症抑制につながることは、日本人でも明らかになっている。

しかし、ハイリスク者選定にはいまだ問題があり、最近発表されたメタアナリシスで、HbA1cのみのスクリーニングの感度が0.49でしかなかった。また、様々な糖尿病リスク因子が同定され、日本人対象のリスクスコアも開発されているが、スコア計算に使われる指標はまちまちである。加えて、一時点の状態からその後数年間の発症リスクをとらえたモデルによるもので、毎年健康診断情報が蓄積していくわが国の産業保健の状況に必ずしも適していない。

わが国の産業保健領域では、特定保健指導の開始以来、労働者に保健指導を行う体制が整った。指導の有効性も明らかになりつつある。しかしながらメタボリックシンドロームを対象とした特定保健指導以外明確な保健指導介入の枠組みはなく、40歳以下の若年層からの介入対象者を選ぶための根拠はない。糖尿病に関しては、様々なリスクが同定されているものの、発症予防介入のための明確な「ハイリスク状態」が確立されていない。血糖値の「ハイリスク」基準に関しても、例えばWHOなどでは110-125 mg/dl（境界型糖尿病）であるいっぽう、アメリカ糖尿病学会では100-125 mg/dl（前糖尿病）としており、一貫していない。

健康な若年のうちから長い経過で繰り返し指導機会のある産業保健現場での介入判断に資する、糖尿病発症モデルの知見は十分ではないと考えた。

医療経済学の費用対効果分析などでは、慢性的な健康状態の動的な推移を、マルコフモデルの「状態と遷移」として表して分析することが行われてきた。この場合の「状態」は検査結果や診断などからはっきり定義できるあるものに対し、隠れマルコフモデルでは観察データからあらかじめ定義できない「潜在状態」を機械学習で同定する。

隠れマルコフモデルは、もともと情報・工学系での音声認識や医学では遺伝子解析で用いられてきた。音声の縦断データに隠れマルコフモデルを適用するように、健康診断の血糖縦断データに適用することで、糖尿病の「ハイリスク状態」を（潜在状態として）同定できる可能性を考えた。

(2) 本研究では、血糖縦断データに隠れマルコフモデルを適用して糖尿病の「ハイリスク状態」の同定を目指す。脂肪肝などの糖尿病リスクの検討（2-1）やハイリスク状態の介入の効果（2-2）なども同時検討する必要があった。

### 2. 研究の目的

(1) 本研究では、隠れマルコフモデルを労働者の健康診断データに適用して、糖尿病発症モデルを構築することを目的とした。モデルは、未病（健康）の状態、ハイリスク状態、糖尿病の状態を決め、状態から状態へ、毎年行き来（遷移）がある動的なものとなるよう目指した。

これによってその年の健康診断結果から、より確実な介入ターゲットとしてのハイリスク状態を同定することができると考えた。また、ハイリスク状態から糖尿病発症予防だけでなく、未病状態からハイリスク状態への遷移に対する予防の知見も得られることで、若年の未病状態から一貫して包括的な糖尿病発症予防が可能になるとも考えた。

加えて、毎年健診データが蓄積していくわが国の産業保健の特徴を活かした、包括的な糖尿病予防介入法の構築が可能になると考えた。

(2) 加えて、本研究のために集めた労働者の健康診断データセットを利用して、労働者の糖尿病予防にまつわる追加の疫学的検討を行った。肥満者における「脂肪肝の糖尿病リスクの検討」（2-1）および「特定保健指導の糖尿病発症予防の検討」（2-2）である。

### 3. 研究の方法

(1) 機械学習による推定にともなう過学習の問題を考慮して、2つの独立した日本の労働者の健康診断データを利用して2型糖尿病発症モデルを構築した。具体的には、A県の健診機関を受診した労働者の健康診断データセット①（ $n = 28,254,113,768$  人年）と、全国の製薬業労働者の健康診断データセット②（ $n = 51,794,179,933$  人年）を利用し、空腹時血糖の縦断データに隠れマルコフモデルを適用した。2型糖尿病の「ハイリスク状態」の知見を得る目的から、モデルは、3状態モデル（正常状態-ハイリスク状態-糖尿病）で、空腹時血糖126 mg/dlで2型糖尿病とするいっぽう、正常状態とハイリスク状態はそれぞれガウス分布データを出力する潜在状態であると仮定した。

(2-1) 「脂肪肝の糖尿病リスクの検討」については、ベースライン時の健康診断の腹部エコーで診断した脂肪肝が、非肥満の労働者の糖尿病リスクにどのように関わるか、生存分析で検討を行った。

(2-2)「特定保健指導の糖尿病発症予防の検討」については、特定保健指導が40歳から開始されることを利用して、特定保健指導の糖尿病発症予防効果を回帰不連続分析で検討した。

#### 4. 研究成果

##### (1)「隠れマルコフモデルによる糖尿病発症モデル」

3状態隠れマルコフモデルをデータセット①の13,768人年データに当てはめると、モデルは「正常状態」が $90.3 \pm 7.6$  mg/dl、「ハイリスク状態」が $104.6 \pm 7.1$  mg/dlであることを示した。このモデルに基づいて、1年後の正常からハイリスクへの移行確率、高リスクから正常への移行確率、そしてハイリスクから糖尿病への移行確率は、それぞれ7.9%、0.01%、6.0%と推定された(図1)。データセット②(179,933人年)での同様検討で、ほぼ同じ結果が再現された。

本研究の隠れマルコフモデルによる「ハイリスク状態」の推定分布( $104.6 \pm 7.1$  mg/dl)から、2型糖尿病のハイリスクを110 mg/dLとしているWHOのカットオフポイントは高すぎる可能性が示唆された。

さらに、「ハイリスク状態」から「正常状態」への移行確率の推定値が0.01%と非常に低いことから、「ハイリスク状態」に至った者は「正常状態」に戻ることはほとんどない可能性が示唆された。したがって2型糖尿病予防のためには、特定保健指導の文脈とは別に「ハイリスク状態」として特定された個人に対する毎年のモニタリングと継続的な介入が必要であること、「ハイリスク状態」に至らないための「正常状態」の者たちへのポピュレーションアプローチが重要になること、以上の2点を考えた。

このように、隠れマルコフモデルにもとづく2型糖尿病発症モデルは、発症の実態と予防戦略についての理解を向上させる可能性があると考ええる。

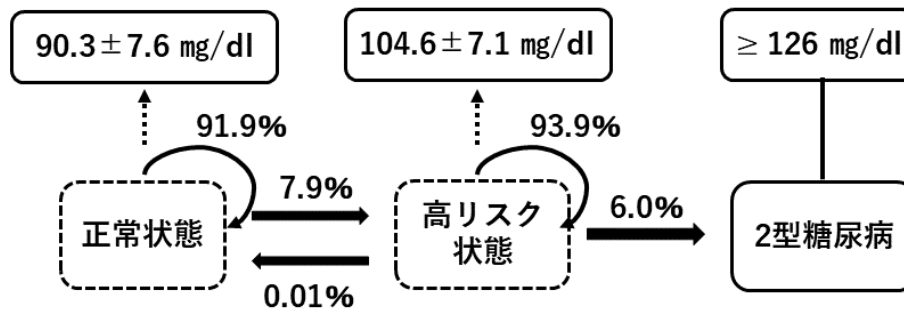


図1. 隠れマルコフモデルにもとづく2型糖尿病発症モデル

##### (2-1)「脂肪肝の糖尿病リスクの検討」

上記(1)で空腹時血糖値の時系列データのみで2型糖尿病発症モデルを構築したが、糖尿病の代表的なリスク因子である脂肪肝と(日本人の特徴である)非肥満の糖尿病について、男女の違いの有無を生存解析で検討した。男性の場合、脂肪肝のないBMIが23-25kg/m<sup>2</sup>(非肥満の場合はDMのリスクはなかったが(HR 1.20、95%CI 0.90~1.60)、脂肪肝のある男性ではDMのリスクがあった(HR 2.35、95%CI 1.77~3.11)。一方、脂肪肝の有無にかかわらず、非肥満の23-25kg/m<sup>2</sup>の女性の両方で、女性のDMリスクが増加していた(それぞれHR 4.43、95%CI 2.21~8.88、HR 1.87、95%CI 1.04~3.35)(図2)。これらの結果は、脂肪肝とDMリスクの組み合わせは性別によって異なり、女性ではなく男性のリスクであることを示唆された。

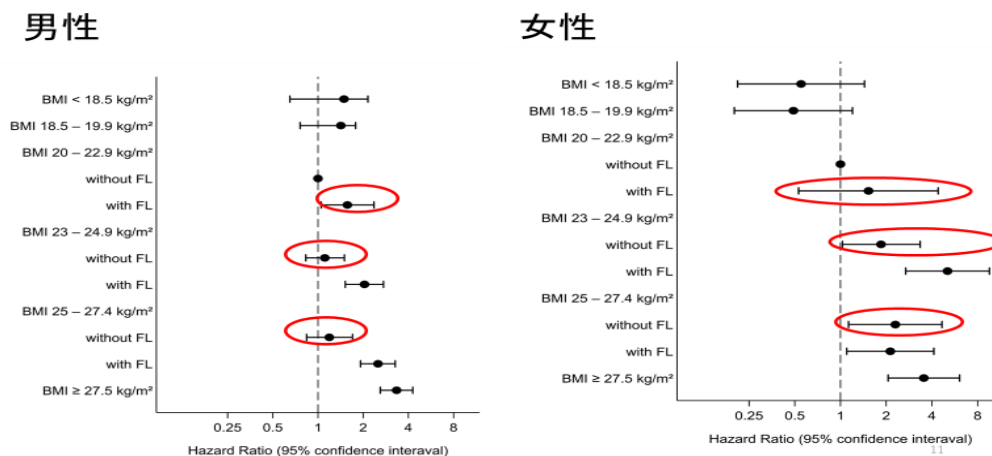


図2. 男女別の脂肪肝とBMIの2型糖尿病発症リスク

(2-2) 「特定保健指導の糖尿病発症予防の検討」

上記(1)では、2型糖尿病の「ハイリスク状態」の同定を目指したが、それは「ハイリスク状態」での介入により糖尿病発症予防ができることが前提となる。メタボリックシンドロームをハイリスク状態とした特定保健指導について糖尿病発症予防を検討した。40歳以上の腹部肥満を伴う空腹時血糖100mg/dl以上の男性で、特定保健指導の介入により10.1減少/1,000人年に相当する糖尿病発症率の減少。相対リスクは0.75と推定された。しかし、腹部肥満がなく介入の対象ではなかった人々の間では、40歳時点での糖尿病罹患率に大きな変化はなかった(図3)。特定保健指導は、肥満のある労働年齢男性の糖尿病発症予防効果示唆された

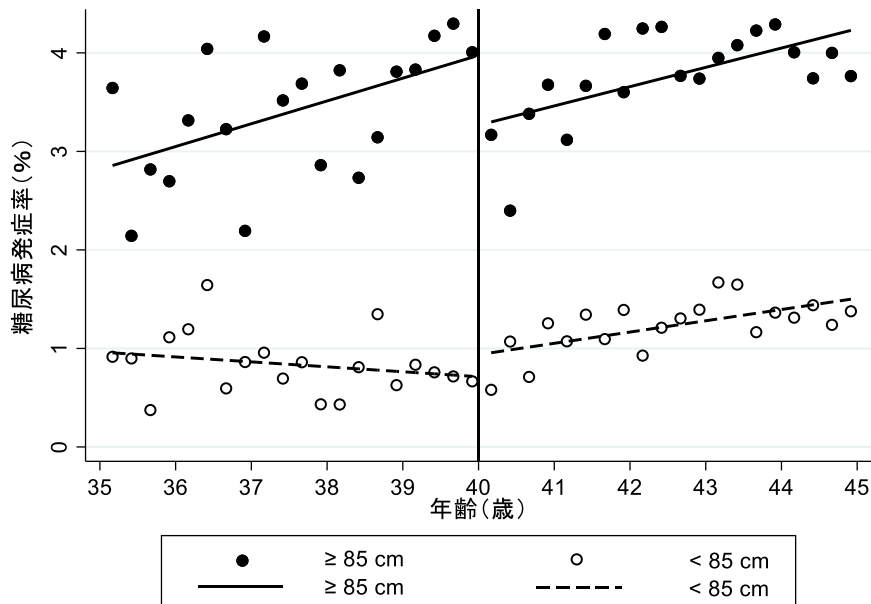


図3. 腹部肥満(85cm以上)の有無別の特定保健指導開始40歳での糖尿病発症率低下

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 3件）

1. 著者名 Narisada Akihiko, Shibata Eiji, Hasegawa Tomomi, Masamura Nobue, Taneda Chitose, Suzuki Kohta	4. 巻 12
2. 論文標題 Sex differences in the association between fatty liver and type2 diabetes incidence in non obese Japanese: A retrospective cohort study	5. 発行年 2021年
3. 雑誌名 Journal of Diabetes Investigation	6. 最初と最後の頁 1480 ~ 1489
掲載論文のDOI (デジタルオブジェクト識別子) 10.1111/jdi.13496	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Narisada Akihiko, Shibata Eiji, Hasegawa Tomomi, Wakayama Rei, Suzuki Kohta	4. 巻 189
2. 論文標題 The impact of the National Health Program on diabetes incidence among working-age men with prediabetes: A regression discontinuity analysis of a nation-wide database in Japan	5. 発行年 2022年
3. 雑誌名 Diabetes Research and Clinical Practice	6. 最初と最後の頁 109946 ~ 109946
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.diabres.2022.109946	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計12件（うち招待講演 0件／うち国際学会 4件）

1. 発表者名 Akihiko Narisada, Eiji Shibata, Rei Wakayama, Kohta Suzuki
2. 発表標題 The Effect of the National Health Program in Japan on the Development of Diabetes among Men with Prediabetes: Regression Discontinuity Evidence from the Minimum Target Age of the Program
3. 学会等名 Society for Epidemiologic Research 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 成定 明彦、柴田 英治、若山 怜、鈴木 孝太
2. 発表標題 特定保健指導による労働者世代男性の糖尿病発症予防効果
3. 学会等名 第94回日本産業衛生学会
4. 発表年 2021年

1. 発表者名 Akihiko Narisada, Kentaro Katahira, Masahiro Matsunaga, Tomohisa Nagata, Koji Mori, Kohta Suzuki
2. 発表標題 Hidden Markov models for the onset of type 2 diabetes in Japanese working population
3. 学会等名 Society for Epidemiologic Research 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 成定明彦、片平健太郎、松永昌宏、永田智久、森晃爾、鈴木孝太
2. 発表標題 隠れマルコフモデルによる糖尿病高リスク状態の推定：コラボヘルス研究
3. 学会等名 第93回日本産業衛生学会
4. 発表年 2020年

1. 発表者名 成定明彦、片平健太郎、松永昌弘、鈴木孝太
2. 発表標題 隠れマルコフモデルによる男性労働者の糖尿病発症モデルの推定：愛知健康増進研究
3. 学会等名 第30回日本疫学会
4. 発表年 2020年

1. 発表者名 成定明彦、柴田英治、鈴木孝太
2. 発表標題 日本人非肥満者における脂肪肝と糖尿病リスクの男女別の検討：愛知健康増進研究
3. 学会等名 第31回日本疫学会
4. 発表年 2021年

1. 発表者名 Akihiko Narisada
2. 発表標題 Effects of Obesity on Progression of Diabetes Status in Young and Old High Risk Japanese Male Workers: Multistate Analysis
3. 学会等名 Society for Epidemiologic Research 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 成定明彦
2. 発表標題 隠れマルコフモデルによる男性労働者の糖尿病高リスク状態の推定：愛知健康増進研究
3. 学会等名 日本産業衛生学会東海地方会
4. 発表年 2019年

1. 発表者名 成定明彦
2. 発表標題 隠れマルコフモデルによる男性労働者の糖尿病発症モデルの推定：愛知健康増進研究
3. 学会等名 第30回日本疫学会
4. 発表年 2020年

1. 発表者名 成定明彦、柴田英治、梅村朋弘、鈴木孝太
2. 発表標題 労働者の糖尿病緊急症の検討
3. 学会等名 日本産業衛生学会東海地方会
4. 発表年 2020年

1. 発表者名 Kohta Suzuki, Rei Wakayama, Keiki Ueda, Akihiko Narisada
2. 発表標題 Factors associated with maintenance of behaviors based on transtheoretical model: Aichi Health Promotion Study
3. 学会等名 The 22nd World Congress of Epidemiology 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 若山 怜、成定明彦、鈴木孝太
2. 発表標題 行動変容ステージと健診結果の傾向スコアを用いた検討：愛知健康増進研究
3. 学会等名 第81回日本公衆衛生学会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	鈴木 孝太  (Suzuki Kohta)		
研究協力者	柴田 英治  (Shibata Eiji)		
研究協力者	長谷川 共美  (Hasegawa Tomomi)		



## 6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	若山 怜  (Wakayama Rei)		
研究協力者	松永 昌宏  (Matsunaga Masahiro)		
研究協力者	片平 健太郎  (Katahira Kentaro)		
研究協力者	梅村 朋弘  (Umemura Tomohiro)		
研究協力者	正村 伸江  (Masamura Nobue)		
研究協力者	千歳 種子田  (Taneda Chitose)		
研究協力者	永田 智久  (Nagata Tomohisa)		
研究協力者	森 晃爾  (Mori Koji)		

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------