様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　７ 年　 ６ 月　 ３ 日現在

機関番号：１４４０１
研究種目：基盤研究(C)（一般）
研究期間：2022～2024
課題番号：２２Ｋ１００７４

研究課題名（和文）長期縦断データに基づいた機械学習による認知機能低下予測モデルの作成

研究課題名（英文）Machine learning models for predicting cognitive decline based on long-term longitudinal data

研究代表者
高橋　利士（Takahashi, Toshihito）

大阪大学・大学院歯学研究科・招へい教員

研究者番号：７０６１０８６４

交付決定額（研究期間全体）：（直接経費）　　3,200,000円

研究成果の概要（和文）：　超高齢社会に突入した我が国において，一人でも多くの高齢者が自立した生活を営むためには，個々の高齢者に対して，要介護状態の原因である疾患の発症やそのリスク因子を予測し，高齢者ごとに必要な対策を講じることが重要であると思われる．
　そこで，要介護の主な原因である認知症に注目し，「口腔内の状態から認知症の発症を予測することができれば，歯科的なアプローチにより介護予防を行うことができる」との仮説のもと，12年にわたる3000名の長期縦断調査のデータを用いて口腔の状態を含む膨大な数の因子から認知機能の低下を機械学習により予測するモデルを作成する．

研究成果の学術的意義や社会的意義
　本研究で作成する認知機能低下予測モデルは，健康長寿の要因を探るため10年以上にわたり，3000名もの学際的な大規模長期縦断調査，老年学研究により得られた身体的因子，社会的因子，心理的因子のデータに，口腔因子（残存歯数や歯周病の状態など）や口腔機能（咬合力や唾液分泌機能など）のデータを加えた包括的なデータに基づいたものである．
　したがって，本研究で作成された予測モデルは，口腔に関するデータ以外も考慮されているだけでなく，これまでにない長期的な予測が可能な信頼性の高い予測モデルであると思われる．

研究成果の概要（英文）：As our country enters a super-aged society, it is crucial to predict the onset of diseases that lead to the need for long-term care, as well as their risk factors, for each elderly individual. By doing so, appropriate measures can be implemented to help as many elderly people as possible maintain independent lives.
　With this in mind, we focused on dementia, which is a major cause of the need for long-term care. Based on the hypothesis that "if the onset of dementia can be predicted from oral conditions, preventive care can be achieved through dental approaches," we aim to develop a machine learning model to predict Cognitive declin. This model will be built using data from a 12-year longitudinal study involving 3,000 participants and analyzing a vast number of factors, including oral health conditions.

研究分野：高齢者歯科学

キーワード：高齢者

様　式　Ｃ－１９、Ｆ－１９－１（共通）

１．研究開始当初の背景

高齢者が要介護状態に陥ることを防ぐ「介護予防」は，個々の高齢者の問題だけでなく，医療費削減，労働力人口への負担軽減という点で，社会全体の問題の解決法としても重要である．そのためには，高齢者が要介護状態となるリスク因子とその重要度を予測し，その人のリスク因子に応じた対策を講じることが必要であると思われる．

現時点ではそれらのリスク因子の中に口腔の状態に関するものは含まれていないが，我々が 2010 年からおこなっている地域在住高齢者を対象とした学際的な大規模コホートの長期縦断調査のデータから，口腔の状態と認知機能の低下との間に関連があることが示唆されている．

２．研究の目的

口腔の状態を因子として含む認知機能の低下（軽度認知障害 Mild Cognitive Impairment: MCI）を予測するモデルを作成することにより，将来的に口腔の状態から個々の高齢者の認知症発症のリスクを予測し，予測結果に基づいて歯科医療関係者が，個々の高齢者に合わせた口腔の管理を行うことで，介護予防が期待できる．

本研究は，身体的因子，社会的因子，心理的因子のデータに，口腔因子（残存歯数や歯周病の状態など）や口腔機能（咬合力や唾液分泌機能など）のデータを加えた包括的なデータから，認知機能低下を機械学習により予測するモデルを作成することを目的とする．

３．研究の方法

(1) 口腔機能の因子を含む因果探索 LinGAM モデルの作成

本研究の対象者は，2012 年および 2013 年に健康長寿研究のベースライン調査に参加した 70 歳と 80 歳の地域在住高齢者の中で，６年後の追跡調査時にも参加した 920 名（70 歳:548 名（男性 271 名，女性 277 名），80 歳:372 名（男性 184 名，女性 188 名））とした．

調査項目は，口腔内因子（残存歯数，歯周病の状態，最大咬合力，口腔関連 QoL(GOHAI)）と，性別，年齢，BMI，教育歴，収縮期血圧，喫煙状況，飲酒状況，精神健康状態(WHO5)，心理状態(GDS5)，認知機能(Moca-J)とした．なお，Moca-J の点数が 22 点未満のものを認知機能が低下していると判断した．

上記の対象者の中で，ベースライン時に認知機能の低下がみられなかった 607 名（70 歳:405 名（男性 188 名，女性 217 名），80 歳:202 名（男性 96 名，女性 106 名）のデータを用いて各年齢と性別ごとの Lingam モデルを作成した．本モデルの作成の際には，調査項目の中で Moca-J の値を目的変数，その他の項目を説明変数とし，すべての変数は連続変数として利用した．また，データ欠損については，補完せずにモデルの作成を行った．

Lingam モデルの作成は，プログラム言語には python3 を使用し，Google Collaboratory の環境下でおこなった．また，各年代，性別ごとの認知機能低下に影響を及ぼす各因子の相対的な特徴量を，lightGBM アルゴリズムを用いて算出した．

(2) 機械学習アルゴリズムを使用した認知機能予測モデルの作成

①認知機能低下予測カテゴリー変数モデルの作成

Lingam モデルと同じ対象者を使用して，様々なアルゴリズムを使用して認知機能を予測するモデルを作成した．データ欠損については，MisForest を用いて補完した．

本モデルでは，前述の調査項目をすべてカテゴリー変数化し，各因子のカットオフ値は以下の通りとした；残存歯数（20 本以上，20 本未満），歯周病の状態（4 ㎜以上の歯周ポケットなし，あり），最大咬合力（200Ｎ以上，200Ｎ未満），性別（男性，女性），年齢（70 歳，80 歳），BMI（25 未満，25 以上），教育歴（9 年以下，10～12 年，13 年以上），収縮期血圧値（140 以上，140 未満），糖尿病（HbA1c 値 6.0 未満，HbA1c 値 6.0 以上），喫煙状況（喫煙あり，なし），飲酒状況（あり，なし），精神健康状態（WHO5 値 13 未満，13 以上），心理状態（GDS5 値 4 以上，3 以下），認知機能（MoCA-J 値 22 以上，22 未満），居住状況（同居者あり，なし），近所付き合い（あり，なし）．

予測モデルの作成は，認知機能の低下予測モデルは，認知機能の低下プログラム言語は Python3 を使用し，Google Collaboratory 環境下でおこなった．モデルの作成には，対象者 607 名をランダムに訓練群（424 名）とテスト群（183 名）に分類し，Artificial Neural Network, Logistic Regression, Support Vector Machine, Naïve Bays, K 近傍法，K 平均法，Gradient Boost, XGBoost, Random Forest, Light GBM の 10 種類のアルゴリズムを使用した．上記のアルゴリズムで作成したそれぞれの予測モデルを，ROC 曲線ならびに AUC 値を使用して評価し，最も AUC が高いモデルを認知機能低下予測モデルとした．その認知機能低下予測カテゴリー変数モデルを使用した場合の混同行列と Accuracy, precision, recall, F-measure を算出した．

②認知機能低下予測連続変数モデルの場合

認知機能低下予測カテゴリー変数モデルと同じ対象者を使用として，カットオフ値によるカ

テゴリー変数化を行わずに，カテゴリー変数と連続変数が混在した状況での予測モデルを作成した.
各因子の変数の状況は下記の通りとした.
カテゴリー変数：歯周病の状態（4 ㎜以上の歯周ポケットなし，あり），性別（男性，女性），年齢（70 歳，80 歳），喫煙状況（喫煙あり，なし），飲酒状況（あり，なし），居住状況（同居者あり，なし），近所付き合い（あり，なし），認知機能（MoCA-J 値 22 以上，22 未満），
連続変数：残存歯数，最大咬合力，教育歴，BMI，収縮期血圧値，糖尿病（HbA1c 値），精神健康状態（WHO5 値），心理状態（GDS5 値）．また，データ欠損の補完については MisForest を用いて行った.
　認知機能低下予測連続変数モデルは, 認知機能の低下を目的変数とし, 上記のカテゴリー変数, 連続変数をそれぞれの説明変数とし，Artificial Neural Network, Logistic Regression, Support Vector Machine, Naïve Bays, K 近傍法，K 平均法，Gradient Boost, XGBoost, Random Forest, Light GBM の 10 種類のアルゴリズムを用いた予測を行い, 最も AUC 値が高いアルゴリズムを使用したモデルを最終的な認知機能低下予測連続変数モデルとした. なお, 訓練群ならびにテスト群の振り分けは認知機能低下予測カテゴリー変数モデル作成の場合と同様とした.

４．研究成果

(1) 口腔機能の因子を含む因果探索 LinGAM モデルの作成
　本研究の対象者の中で，６年後に認知機能が低下したものは 155 名(25.5%)であった. 各性別, 年齢群における因果関係は下記のようになった.
①70 歳男性群の場合
　口腔内因子の中では残存歯数が口腔関連 QoL（GOHAI），心理状態（GDS5）を通じて認知機能の低下に間接的な因果関係がある結果となった. 口腔内因子以外では, 教育歴（education year），喫煙状況（Smoke）が認知機能の低下に直接的な因果関係がある結果となった（図 1）.

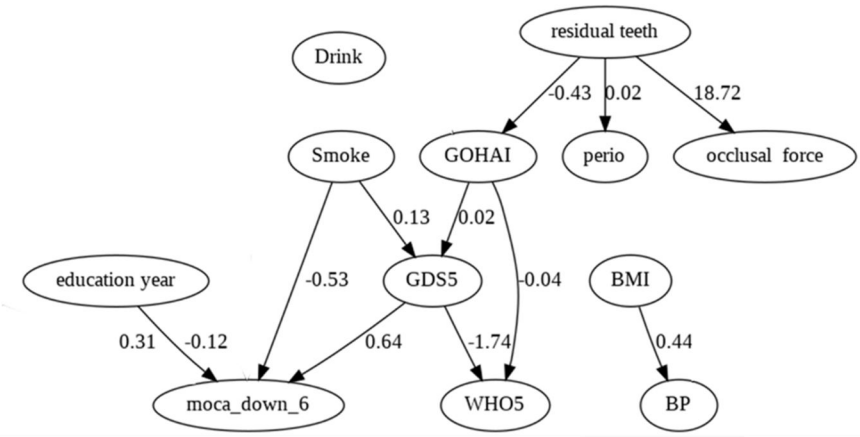

図 1．70 歳男性群における LinGAM モデル

②70 歳女性群の場合
　口腔内因子の中では, 残存歯数（residual teeth）が認知機能低下に直接的な因果関係があり, 口腔関連 QoL（GOHAI）が BMI を通じて間接的な因果関係がある結果となった．口腔内因子以外では，心理状態（GDS5）が精神健康状態（WHO5）を通じて，教育歴（education year），喫煙状況（Smoke），飲酒状況（Drink）が口腔関連 QoL（GOHAI）を通じて，認知機能の低下に間接的な因果関係がある結果となった（図 2）.

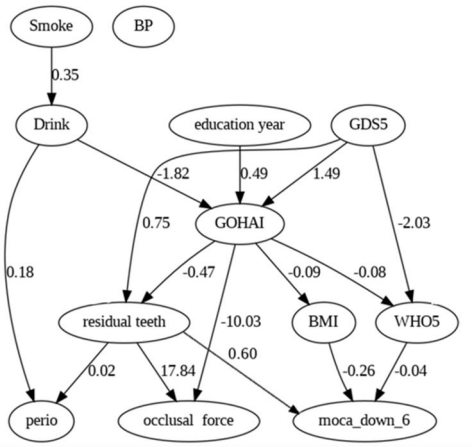

図 2．70 歳女性群における LinGAM モデル

③80 歳男性群の場合

口腔内因子，口腔内因子以外のいずれにおいても，認知機能低下に直接的，あるいは間接的な因果関係がある因子はなかった（図 3）.

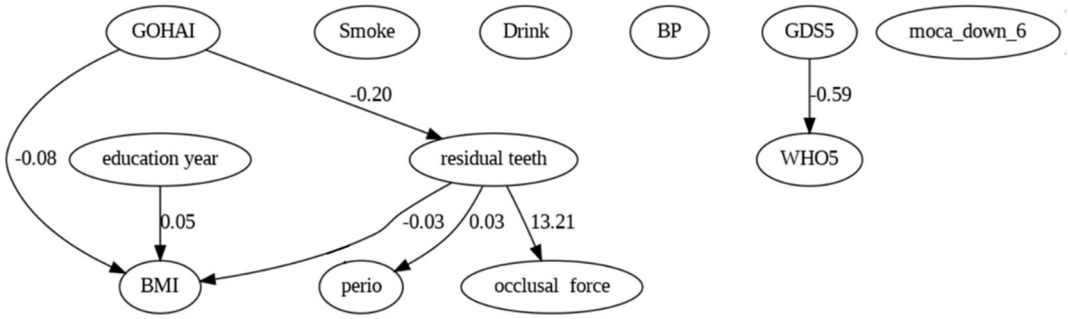

図 3．80 歳男性群における LinGAM モデル

④80 歳女性群の場合

口腔内因子の中では，残存歯数（residual teeth）**が認知機能低下に直接的な因果関係がある結果となった. 口腔内因子以外においては,** BMI **と血圧**（BP）**が直接的に，教育歴**（education year）**が精神健康状態**（WHO5）**を通じて間接的に認知機能低下に因果関係がある結果となった**（**図４**）**.**

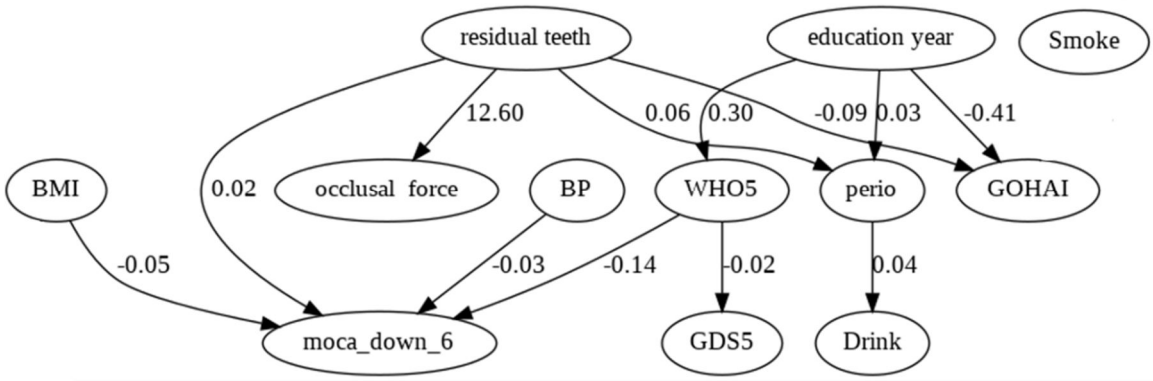

図 4．80 歳女性群における LinGAM モデル

(2) 機械学習アルゴリズムを使用した認知機能予測モデルの作成

①認知機能低下予測カテゴリー変数モデルの作成

各アルゴリズムにより作成したモデルの ROC **曲線ならびに** AUC **値を図** 5 **に示す．**AUC **は最も大きいアルゴリズムは** Random Forest（AUC＝0.712）**であった.**

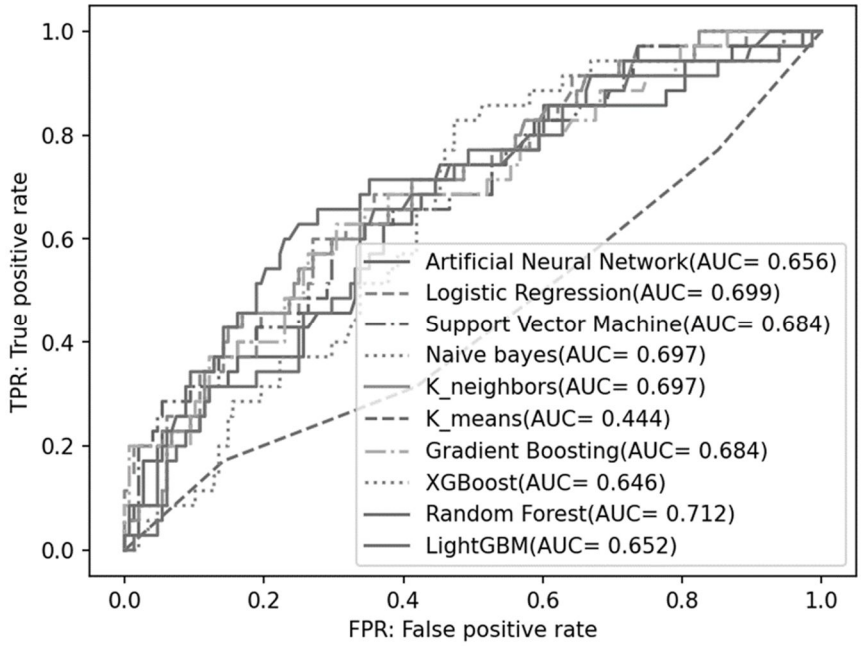

図 5．認知機能低下予測カテゴリー変数モデルの ROC 曲線

②認知機能低下予測連続変数モデルの場合

各アルゴリズムにより作成したモデルの ROC 曲線ならびに AUC 値を図 6 に示す．AUC 値が最も大きいモデルはロジスティック回帰分析を使用した場合（AUC＝0.691）であった．

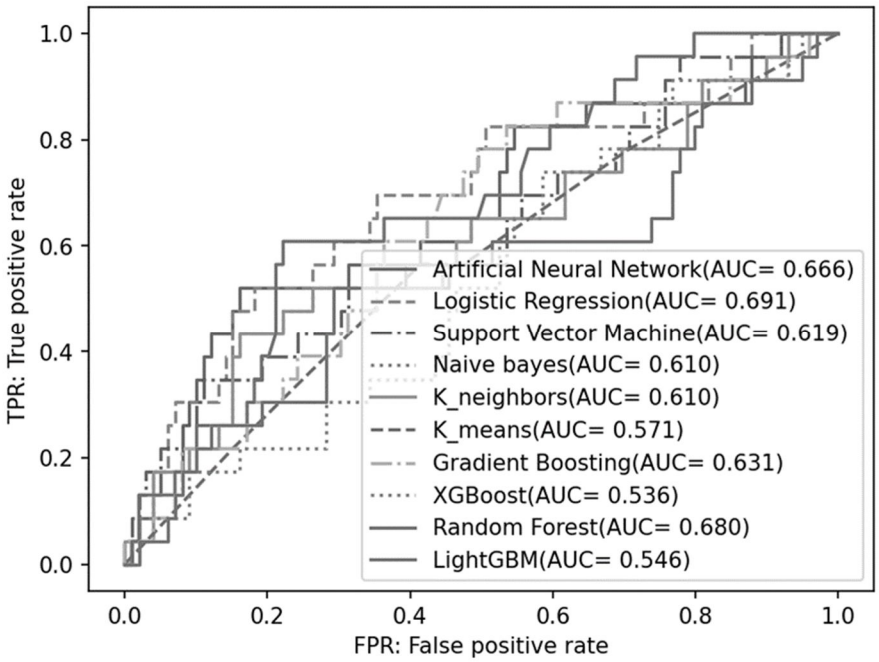

図 6．認知機能低下予測連続変数モデルの ROC 曲線

70 歳，80 歳の地域在住高齢者を対象として，６年後の認知機能低下予測モデルの製作を試みた結果， Lingam **モデルにより，高齢者の認知機能低下と残存歯数が間接的，あるいは直接的に因果関係があることが示唆された．**

また，機械学習を用いた認知機能低下予測モデルを，各説明変数において閾値を設定してカテゴリー変数とした場合と，閾値を設定せず連続変数とした場合の２種類を作成した結果，カテゴリー変数とした予測モデルの方が，わずかではあるが認知機能予測モデルとしての精度が高くなることが分かった．しかしながら，AUC の値は 0.712 にとどまり，モデル自体の精度としては低くとどまった．

この原因として，説明変数の数に対して全体の数が少なかったこと，データに欠損値が多く，補完した値の数が多かったこと，などが考えられる．さらに予測モデルの精度を上げるためにはデータ数を増やすことやデータ整理を行う必要があると思われる．

５．主な発表論文等

〔雑誌論文〕　計2件（うち査読付論文　2件／うち国際共著　1件／うちオープンアクセス　2件）

１．著者名 Takahashi T, Hatta K, Ikebe K	４．巻 59
２．論文標題 Risk factors of cognitive impairment: Impact of decline in oral function	５．発行年 2023年
３．雑誌名 Jpn Dent Sci Rev	６．最初と最後の頁 203-208
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jdsr.2023.06.006.	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 Mameno T, Moynihan P, Nakagawa T, Inagaki H, Akema S, Murotani Y, Takeuchi S, Kimura A, Okada Y, Tsujioka Y, Higashi K, Hagino H, Mihara Y, Kosaka T, Takahashi T, et al.	４．巻 145
２．論文標題 Exploring the association between number of teeth, food intake, and cognitive function: A 9-year longitudinal study	５．発行年 2024年
３．雑誌名 Journal of Dentistry	６．最初と最後の頁 104991～104991
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jdent.2024.104991	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 該当する

〔学会発表〕　計1件（うち招待講演　0件／うち国際学会　1件）

１．発表者名 Higashi K, Takahashi T, Mameno T, Nozaki K, Hatta K, Murotani Y, Tsujioka Y, Akema S, Seto E, Okada Y, Takeuchi S, Mihara Y, Wada M, Maeda Y, Ikebe K
２．発表標題 Causal relationships between cognitive decline and oral factors revealed by structural equation models
３．学会等名 ECG Annual Congress Stockholm 2023（国際学会）
４．発表年 2023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
研究分担者	野崎　一徳 (Nozaki Kazunori) (40379110)	大阪大学・歯学部附属病院・准教授 (14401)	

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
研究分担者	豆野　智昭 (Mameno Tomoaki) (50845922)	大阪大学・大学院歯学研究科・助教 (14401)	
研究分担者	八田　昂大 (Hatta Koudai) (60845949)	大阪大学・大学院歯学研究科・招へい教員 (14401)	

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