

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 15 日現在

機関番号：20101

研究種目：基盤研究(C) (一般)

研究期間：2013～2016

課題番号：25460768

研究課題名(和文) 特定健診における動脈硬化性疾患ハイリスク者抽出効率のCART応用による改善の試み

研究課題名(英文) Establishment of classification and regression tree model for assessing of a risk for future occurrence of life-style related disease using a community-based cohort data.

研究代表者

大西 浩文(Ohnishi, Hirofumi)

札幌医科大学・医学部・准教授

研究者番号：20359996

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：地域一般住民健診受診者コホートのデータに基づき、将来の高血圧、糖尿病、心血管疾患罹患を予測する決定木モデル(CART)の作成を行った。高血圧の予測モデルにおいては感度68.2%、特異度71.1%であり、メタボリックシンドロームと予備群を除外した対象での検討においては感度65.3%、特異度84.2%となった。糖尿病の予測モデルにおいては感度54.2%、特異度96.3%であり、心血管疾患の予測モデルにおいては感度7.4%、特異度99.3%であった。決定木モデルを作成して健診受診者集団から将来の生活習慣病罹患ハイリスク者を効率的に抽出できる可能性が考えられた。

研究成果の概要(英文)：We tried to establish a useful Classification and Regression Tree (CART) model for assessing a risk for future occurrence of life-style related disease, which includes hypertension, type 2 diabetes and cardiovascular disease (CVD), using a community-based cohort data. In the predicting model for hypertension, sensitivity and specificity were 68.2% and 71.1%, respectively. After exclusion of individuals with metabolic syndrome (MetS) or pre-MetS, sensitivity and specificity changed into 65.3% and 84.2%, respectively. In the predicting models for type 2 diabetes, sensitivity and specificity were 54.2% and 96.3%, respectively. In the predicting model for CVD, sensitivity and specificity were 7.4% and 99.3%, respectively. To find individuals at high risk for future occurrence of life-style related disease, using a CART model may be useful in the situation of health checkup and health guidance.

研究分野：疫学

キーワード：特定健診・特定保健指導 生活習慣病罹患ハイリスク者 決定木分析 CART 地域一般住民コホート

1. 研究開始当初の背景

わが国の死因の3分の1を占める心筋梗塞や脳卒中などの動脈硬化性疾患を予防することは、日本人の生命予後を延伸するのみならず、進行する高齢社会では生活の質(QOL)や日常生活活動(ADL)レベルの維持をする上でも重要な課題である。また、動脈硬化性疾患を予防するためには、その危険因子となる高血圧、糖尿病、脂質異常症などの種々の生活習慣病自体を予防することは一次予防の観点から極めて重要となる。

近年、内臓脂肪の蓄積を背景として軽度の危険因子が集積することにより、心血管疾患発症や糖尿病発症の強いリスクとなる病態がメタボリックシンドローム(MetS)として世界的にも注目される中、医療制度改革の一環としてこれまでの健診・保健指導の見直しが図られ、平成20年4月からは特定健診・保健指導が施行されている。その中でMetSは重要な骨子として取り上げられており、MetSに該当する者やその予備群を健診により抽出して積極的に保健指導を行うことになっている。抽出ステップではまず腹部肥満やBody Mass Index(BMI)による肥満の有無で対象者が振り分けられ、肥満該当者においては血圧レベル、血糖値、脂質値、喫煙の危険因子の保有数がカウントされ、予備群に該当した場合は動機づけ支援が、MetSに該当する場合は積極支援が行われる。食生活や運動習慣の見直しによって減量を測り、一元的に危険因子を管理して、生活習慣病やその先にある動脈硬化性疾患の予防を行うことになる。このステップの流れは、肥満者の中でも特に減量することにより効果的に血圧や血糖、脂質値の改善が見込まれる対象を抽出することができ、減量を目標においた保健指導を全国どこにいても質の格差なく受けられることを可能にするという点では効果的な健診と考えられる。しかし一方で、最初のステップで肥満に該当しない対象者は、血圧や血糖、脂質値などが受診勧奨レベルに該当しなければ一律情報提供レベルとなり、積極的な支援を受けることはない。さらにいくつかの研究により、非肥満者でも複数の危険因子を保有する者では動脈硬化性疾患罹患リスクが高まることが報告され、現在の特定健診のステップでは情報提供レベルにとどまる非肥満者での動脈硬化性疾患ハイリスク者を見逃す可能性も指摘されている。よって、特定保健指導対象者抽出のステップとして肥満該当者からのハイリスク者抽出の流れに加えて、肥満非該当者の中からハイリスク者を抽出する流れを組み込んでいく必要性が考えられる。

Classification and Regression Tree(CART)は、Breimanらによって1984年に開発された2進木成長予測モデルである。CARTでは予測因子候補の中から適切な予測因子を選択し、分割値で二分岐することでClassification Tree(決定木)を構築し事象

の予測を行う。予測因子の選択及び分割値の算出に際しては、集団のバラツキの指標であるGINIインデックスを使用し、分岐毎にその集団のバラツキが最小になるようにモデルを構築する。工業や気象予報の分野において応用され広く使用されているが、近年、臨床研究の分野においてもCARTを用いた予後予測などの報告がされている。健診受診者のデータベースにCARTアルゴリズムを応用し、将来の高血圧・糖尿病・動脈硬化性疾患の罹患を判別する決定木を作成することによって、肥満非該当者も含めたハイリスク者抽出ステップを検討し、効果的な健診システムの構築に役立てることが可能になると考えられる。

2. 研究の目的

本研究は、継続中の地域一般住民コホ-ト研究のデータを用い、断面および縦断的検討により、メタボリックシンドロ-ムやその予備群に該当する者に加えて特定健診・特定保健指導においては情報提供レベルにとどまる肥満非該当者においても、通常の健診項目に加えて炎症性マ-カ-やインスリン抵抗性などのリスクマ-カ-を組み合わせ、CARTという決定木アルゴリズムを応用することで積極的な介入が必要と考えられる高血圧や糖尿病などの生活習慣病および虚血性心疾患や脳卒中などの動脈硬化性疾患の罹患に対するハイリスク者の抽出効率を改善することが可能となるかを検討することを目的とする。

3. 研究の方法

対象は、現在継続中の地域住民コホ-ト研究である端野・壮瞥町研究の参加者である。新たに収集する健診データと既存のデータを統合し、端野・壮瞥町の約3000人のコホ-トのデータベースを構築し、CARTアルゴリズムを用いて既に観察済みの高血圧、糖尿病、動脈硬化性疾患イベント発症の有無を効率良く判別する決定木を作成した。特定健診において、動機づけ支援や積極支援の対象となる者を除外した場合の分岐要因の差異についても検討を行った。1994年には腹囲径の測定を行っていることからMetSの判定を行う初年度データとしてCARTモデルを作成し、2002年を初年度とする検証コホ-トを設定して作成したCARTモデルの妥当性について検討を行った。

決定木分析の分岐要因として用いる項目としては、身長、体重、腹囲径、BMI、収縮期血圧値(SBP)、拡張期血圧値(DBP)、空腹時血糖値(FPG)、脂質値などの通常の健診項目に加えて炎症マ-カ-である高感度CRPやインスリン抵抗性の指標であるHOMA-IRなどのマ-カ-を測定することによる判定効率の向上についても検討した。また、各分岐要因のカットオフ値と現在の特定健診における保健指導判定基準とを比較することで、分岐カ

ットオフ値の有用性についての検討を行った。

4. 研究成果

平成 25 年度はこれまで蓄積している健診データを用いて、将来の高血圧罹患者を効率良く予測するモデルについて CART 法を用いて検討した。1994 年の健診受診者を 2007 年まで追跡したデータベースを用いたこれまでの検討で、加齢、肥満、高血圧家族歴、空腹時血糖高値、 β -GTP 高値が高血圧罹患の危険因子であった。年齢 50 歳、BMI 25、高血圧家族歴あり、空腹時血糖値 100mg/dl、 β -GTP が男性 25IU/L、女性 16IU/L の保有数が多い対象ほど将来の高血圧罹患リスクが高いことも確認できた。よってこれらの 5 項目に正常高値血圧（収縮期血圧 130-139mmHg かつまたは拡張期血圧 85-89mmHg）を加えて、将来の高血圧罹患者を判別する CART 法によるモデルの作成を行った。年齢、50 歳以上、正常高値の有無、 β -GTP 高値、血糖値高値の順で分けることにより、将来の高血圧罹患者を感度 68.2%、特異度 71.1%で判別できた。

平成 26 年度は、作成した高血圧予測 CART モデルの構成要素や寄与度、各構成要素のカットオフ値の検討、および初年度の異なる集団における妥当性の検証を行った。高血圧罹患予測決定木分析を男女別に行うと、男性では年齢、血圧値、 β -GTP 高値、血糖高値の順で、女性においては血圧値、年齢、 β -GTP 高値、血糖高値の順で関与しており、男性で感度 69.2%、特異度 67.0%、女性で感度 73.7%、特異度 70.2%で判別可能であり、男女で各項目の寄与に若干の順序の違いはあるものの、同様の構成要素で判別できることが確認された。また上記の高血圧予測 CART モデルは 1994 年を初年度とした集団での検討であるが、2002 年を初年度とする集団に本モデルを使用した場合の妥当性を検討すると、感度 72.2%、特異度 35.6%となり、特異度が低下することが確認された。CART による各カットオフ値は、年齢が 45.5 歳、 β -GTP が男性 24.5、女性 16.5、空腹時血糖値が男性 90.5、女性 99.5、収縮期血圧が男性で 121.5、女性 129.5 といった値が導かれており、保健指導の現場で使用する判定値に近い数値となり得る可能性が考えられた。

平成 27 年度は、特定健診・特定保健指導の場での使用を考慮し、メタボリックシンドローム該当および予備群該当者は面談や保健指導対象であり、それら該当者を除外した集団における高血圧発症ハイリスク者の決定木を検討した。1994 年の男女合わせた対象では、年齢、収縮期血圧（SBP）、喫煙、FPG、BMI、トリグリセリド（TG）、 β -GTP の要素からなる CART モデルにて、感度 65.3%、特異度 84.2%で予測できた。男女別に検討を行うと、男性では年齢、SBP、 β -GTP、FPG、からなる CART モデルにて感度 75.0%、特異度 73.3%で

予測可能となり、女性では SBP、age、BMI、喫煙、総コレステロール（TC）、TG、FPG、 β -GTP からなるモデルにて感度 80.3%、特異度 81.2%で予測可能であった。特定健診・特定保健指導の場では面談や保健指導の対象とならない者においても、今回作成したモデルによって高血圧発症に該当する者については積極的な高血圧予防介入を行うことが重要と考えられる。さらに 2002 年を初年度とするデータにおいて、高感度 CRP や空腹時インスリン値、アディポネクチンも候補要因として加えて CART モデルの作成を試みたが、SBP、年齢、BMI、アディポネクチン、インスリン値、FPG からなる CART モデルで感度 12.3%、特異度 98.9%と感度が非常に低い結果となり、新しいマーカーを加えた CART モデルの有用性についてはさらなる検討が必要と考えられた。また分岐候補要因に加えて、追跡期間の短さが感度・特異度に影響する可能性も考えられた。

平成 28 年度は、糖尿病と心血管疾患（CVD）をアウトカムとした決定木分析を行った。糖尿病発症をアウトカムとした CART モデルにおいては 1994 年の男女合わせた対象で、年齢、性別、SBP、喫煙、飲酒歴、FPG、BMI、総コレステロール（TC）、TG、HDL コレステロール、 β -GTP、糖尿病家族歴の候補要素とすると、感度 54.2%、特異度 96.3%で予測の予測能となった。その際、枝分かれに使用された要因とカットオフ値は、順に FPG107.5、BMI>29.8、SBP>119.5、BMI>25.4、年齢 > 48.5 歳、FPG>91.5、BMI>26.0 であった。CVD 発症をアウトカムとした CART モデルの分析においては、同様に 1994 年の男女合わせた対象で、年齢、性別、腹囲径、eGFR、SBP、メタボリックシンドローム、FPG、TG、HDL、TC、喫煙を候補要素とすると、感度 7.4%、特異度 99.3%という結果であった。枝分かれに使用された要因とカットオフ値は、年齢>62 歳、SBP>151、HDL<48、TC>275、eGFR<68、TC>212 であった。感度が低く特異度が高い結果となった理由の一つとして、CVD が高血圧や糖尿病と比較すると罹患率、有病率が低いことが結果に影響した可能性が考えられた。

一般的にこうした予測モデルに関しては、作成コホートでの当てはまりはよいものの、検証コホートでは感度・特異度が落ちる傾向にある。今回の検討においても同様の結果となっており、世代の違いの影響も含めて複数の初年度を設定してモデルを作成して有用性を検討することやより大規模なコホートにおいて同様の検討を行うことも必要であると考えられた。

今回の検討により、高血圧、糖尿病の罹患を予測する CART モデルは比較的効率的にハイリスク者の抽出を行うことが可能になると考えられた。ただし複数の集団が作成されるため、特定健診の現場で使用する場合には、いくつかのグループにまとめて保健指導の対象とするなどの取り扱いも必要となり、さ

らなる検討が必要であると考えられた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 1件)

Ohnishi H, Saitoh S, Akasaka H, Furukawa T, Mori M, Miura T. Impact of longitudinal status change in metabolic syndrome defined by two different criteria on new onset of type 2 diabetes in a general Japanese population: the Tanno-Sobetsu Study. Diabetol Metab Syndr (2016) 8:64 DOI 10.1186/s13098-016-0182-0 (査読有り)

〔学会発表〕(計 3件)

大西浩文, 齋藤重幸, 赤坂 憲, 古川哲章, 森 満, 三浦哲嗣. 地域一般住民における尿中 Na/K 比の変化量と血圧の変化量との関連 端野・壮瞥町研究 . 第 39 回日本高血圧学会総会, 仙台国際センター(宮城県仙台市), 2016. 9. 30-10. 2.

大西浩文. ジョイントセッション2「予防学会員から心リハ学会員へのメッセージ」端野・壮瞥町研究からみた地域一般住民における高血圧罹患リスク要因. 第 20 回日本心臓リハビリテーション学会学術集会, 京都市勧業館みやこめっせ(京都府京都市), 2014. 7. 19-20.

大西浩文, 齋藤重幸, 森 満, 三浦哲嗣. シンポジウム「日本の高血圧疫学研究: 最新の知見」北海道: 端野・壮瞥町研究. 第 36 回日本高血圧学会総会, 大阪国際会議場(大阪府大阪市), 2013. 10. 24-26.

〔図書〕(計 0件)

〔産業財産権〕

出願状況(計 0件)

取得状況(計 0件)

〔その他〕

ホ-ムペ-ジ等

6. 研究組織

(1) 研究代表者

大西 浩文 (OHNISHI Hirofumi)

札幌医科大学・医学部・准教授

研究者番号: 20359996

(2) 研究分担者

齋藤 重幸 (SAITOH Shigeyuki)

札幌医科大学・保健医療学部・教授

研究者番号: 60253994

三浦 哲嗣 (MIURA Tetsuji)

札幌医科大学・医学部・教授

研究者番号: 90199951

森 満 (MORI Mitsuru)

札幌医科大学・医学部・教授

研究者番号: 50175634