

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-64087
(P2012-64087A)

(43) 公開日 平成24年3月29日(2012.3.29)

(51) Int.Cl.	F I	テーマコード (参考)
G06Q 50/22 (2012.01)	G06F 17/60 126G	4C117
A61B 5/00 (2006.01)	A61B 5/00 G	
G06N 3/00 (2006.01)	G06N 3/00 550Z	

審査請求 未請求 請求項の数 9 O L (全 14 頁)

(21) 出願番号	特願2010-208988 (P2010-208988)	(71) 出願人	899000079
(22) 出願日	平成22年9月17日 (2010.9.17)		学校法人慶應義塾 東京都港区三田2丁目15番45号
		(74) 代理人	100118821 弁理士 祖父江 栄一
		(74) 代理人	100102635 弁理士 浅見 保男
		(74) 代理人	100106459 弁理士 高橋 英生
		(74) 代理人	100105500 弁理士 武山 吉孝
		(74) 代理人	100103735 弁理士 鈴木 隆盛

最終頁に続く

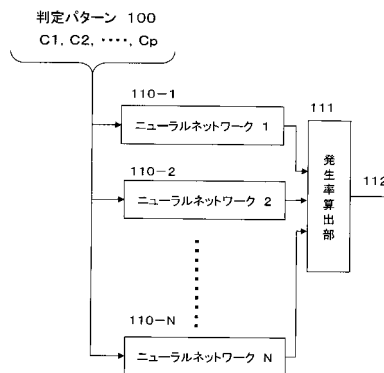
(54) 【発明の名称】 生活習慣病の診断予測装置、生活習慣病の診断予測方法及びプログラム

(57) 【要約】

【課題】 人工ニューラルネットワークを用いて数年の後に生活習慣病を発症する可能性が有るか否かを予測する。

【解決手段】 入力層と少なくとも一層の中間層と出力層を有する三層ネットワークからなるニューラルネットワークを少なくとも年齢、体格指数(BMI)、拡張期血圧(DBP)、HDLコレステロール、LDLコレステロール、HOMA-1Rの6つの項目からなる学習用のパターンによって生活習慣病の発症の可能性を学習させ、健康診断等により得られた診療データから数年の後に生活習慣病を発症する可能性が有るか否かを予測する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

診療データから生活習慣病の発症の可能性を予測するニューラルネットワークを備えた診断予測装置であって、

前記装置は入力層と少なくとも一層の中間層と出力層を有するニューラルネットワークを備え、

前記ニューラルネットワークは、少なくとも年齢、体格指数(BMI)、拡張期血圧(DBP)、HDLコレステロール、LDLコレステロール、インスリン抵抗性指数(HOMA-IR)の6つの項目からなる学習用のパターンによって生活習慣病の発症の可能性を学習したものであり、

前記装置は、前記学習用のパターンの項目に対応するデータを含む診療データを入力する入力手段を備え、

前記入力手段から入力された診療データから生活習慣病の発症の可能性を予測することを特徴とするニューラルネットワークを備えた診断予測装置。

10

【請求項 2】

請求項1のニューラルネットワークを備えた診断予測装置であって、

学習項目は更に収縮期血圧、血漿グルコース、中性脂肪、肝機能(AST, IU/L)、肝機能(ALT, IU/L)、HMW-アディポネクチン、総アディポネクチン、グリコアルブミン、総コレステロール、遊離脂肪酸、インスリン、喫煙の有無を示すデータの少なくとも一つを含むニューラルネットワークを備えた診断予測装置。

20

【請求項 3】

請求項1または請求項2記載のニューラルネットワークを備えた診断予測装置であって、前記生活習慣病はメタボリックシンドローム、耐糖能異常、高血圧、脂質異常症の1つであるであることを特徴とするニューラルネットワークを備えた診断予測装置。

20

【請求項 4】

請求項1または請求項2記載のニューラルネットワークを備えた診断予測装置であって、学習項目は更に高分子量アディポネクチン、総アディポネクチン、インスリン、グリコアルブミン、および遊離脂肪酸を含み、

前記装置は、診療データにおける前記学習用のパターンの項目に対応するデータの入力手段と、前記入力手段から入力された診療データからメタボリックシンドロームの発症の可能性を推定することを特徴とするニューラルネットワークを備えた診断予測装置。

30

【請求項 5】

請求項1または請求項2記載のニューラルネットワークを備えた診断予測装置であって、

学習項目は更に血漿グルコースを含み、

前記装置は、診療データにおける前記学習用のパターンの項目に対応するデータの入力手段と、前記入力手段から入力された診療データから耐糖能異常の発症の可能性を推定することを特徴とするニューラルネットワークを備えた診断予測装置。

【請求項 6】

請求項1または請求項2記載のニューラルネットワークを備えた診断予測装置であって、

学習項目は更に血清中性脂肪を含み、

前記装置は、診療データにおける前記学習用のパターンの項目に対応するデータの入力手段と、前記入力手段から入力された診療データから脂質異常症の発症の可能性を推定することを特徴とするニューラルネットワークを備えた診断予測装置。

40

【請求項 7】

請求項1または請求項6記載のニューラルネットワークを備えた診断予測装置であって、

前記ニューラルネットワークは構成が異なる複数のニューラルネットワークからなり、

前記複数のニューラルネットワークは同一の学習用のパターンによって生活習慣病の発症の可能性を学習したものであり、

前記入力手段から入力された診療データを前記複数のニューラルネットワークに入力し、前記複数のニューラルネットワークの予測結果に基づいて生活習慣病の発症確率を予測することを特徴とするニューラルネットワークを備えた診断予測装置。

50

【請求項 8】

入力層と少なくとも一層の中間層と出力層を有するニューラルネットワークを用いて診療データから生活習慣病の発症の可能性を予測するニューラルネットワークの制御方法であって、

前記方法は、

前記ニューラルネットワークを学習用のパターンにより学習するステップと、

前記学習するステップにより学習されたニューラルネットワークに診療データを入力して生活習慣病の発症の可能性を予測するステップからなり、

前記学習用のパターン及び上記診療データは少なくとも年齢、体格指数(BMI)、拡張期血圧(DBP)、HDLコレステロール、LDLコレステロール、HOMA-IRの6つの項目を含むことを特徴とするニューラルネットワークの制御方法。

10

【請求項 9】

コンピュータに入力層と少なくとも一層の中間層と出力層を有するニューラルネットワークの機能を実現するステップと、

前記ニューラルネットワークを学習用のパターンにより学習するステップと、

前記学習するステップにより学習されたニューラルネットワークに診療データを入力して生活習慣病の発症の可能性を予測するステップからなり、

前記学習用のパターン及び上記診療データは少なくとも年齢、体格指数(BMI)、拡張期血圧(DBP)、HDLコレステロール、LDLコレステロール、HOMA-IRの6つの項目を含むことを特徴とするコンピュータプログラム。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、医療における診断の補助装置とその制御方法及びプログラムに関するものであり、非線形予測技術、特にニューラルネットワークを用いて、健診者の健診データ等から生活習慣病に罹患する可能性を診断し予測する装置と方法及びプログラムに関するものである。

【0002】

病気の診断データあるいはプラントの測定データから病名あるいは故障の原因を推測する、あるいは今後の推移を予測する手段としてニューラルネットワークが利用されている。ニューラルネットワークは複数のニューロン回路により構成される。図12(a)はニューロン回路の例示である。ニューロン回路は多入力(X_1, X_2, \dots, X_n) 1出力(y)の回路であり、入力111(X_1, X_2, \dots, X_n)の各々には重み係数113(W_1, W_2, \dots, W_n)が対応づけられており、重み演算部112により、入力(X_i)と重み係数(W_i)の積が求められる。各演算部112の演算結果は加算回路114により加算され、シグモイド関数、ヘビサイド関数等の閾値関数に従う正規化手段115により正規化され、出力116となる。

30

【0003】

図12(b)はニューラルネットワークの学習を示す図である。例えば「風邪」と認識させたい症状のパターンと「風邪」と認識させたくない症状のパターンを学習パターン161「 A_1, A_2, \dots, A_p 」として多数用意する。また、各学習パターンには正解の出力、即ち「風邪」あるいは「風邪ではない」を表す「 A_2 」が正解データ162として付加されている。

40

学習を開始する前のニューロン回路の各重み係数(W_i)は所定の値に初期化されている。学習パターン161の「 A_1 」をニューラルネットワークの入力部 X_1 に、「 A_2 」を入力部 X_2 と、学習パターン161に各要素をニューラルネットワークの対応する入力部に入力する。入力された学習パターン161の各要素は中間層152と出力層154のニューロン回路により順次演算され、出力端155に出力「 Z 」を出力する。

【0004】

比較回路163は出力端155からの出力「 Z 」と学習パターン161に付加されている正解データ「 A_2 」と比較する。重み変更手段164は出力「 Z 」と正解データ「 A_2 」が等しくなるよ

50

うに、予め定められた方法に従って各重み係数(W_i)の値を変更する。準備した全ての学習パターンに対して正しい値を出力する重み係数(W_i)を決定して学習を終了する。学習が終了したニューラルネットワークに未知の症状のパターンが入力されると、当該症状のパターン要素を各ニューロン回路の重み係数に従って演算し、「風邪」に該当するか否かを示す情報を「1」あるいは「0」として出力層154から出力する。ニューラルネットワークとその学習に関する公知文献には以下のものが存在する。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2005 - 319301号公報

10

【特許文献2】特開2005 - 293241号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

推論回路には多重ロジスティック回帰(MLR)によるものと、人工ニューラルネットワーク(Artificial Neural Network : ANN)によるものが知られている。MLRによる分析は重要な項目(因子)を検知する強力な方法である。しかし、MLRは原因と結果の間に線形の関係(ロジスティック回帰)が必要であり、原因と結果の間に非線形の関係が存在する場合、MLRは原因から正しい結果を推論することが困難である。

人工ニューラルネットワーク(ANN)システムは人間の脳の処理能力をシミュレートするものであり、入力と出力の間に存在する関連性が学習されると、与えられた未知の入力から正しい出力を予測することが可能である。

20

ANNは予測値を求める時に非線形の表現を使用するため、MLRに存在する非線形の問題の悪影響を受けず、個々の入力に対して合理的な予測を提供することが可能である。ANNはいくつかの分野でMLRより良い予測結果を出力することが知られており、近年ANNを医療分野で使用することが注目されている。例えば、下部消化管出血の再発、経皮的内視鏡胃瘻造設術後の結果、慢性C型肝炎の治療効果の予測において古典的な線形の推定方法より優れた結果を出している。

【0007】

日本において、西洋化された食生活の導入と身体活動の減少により特徴付けられた生活様式の広まりに伴い、腹部肥満、2型糖尿病、高血圧症、脂質異常症、および動脈硬化性病変のような生活習慣病が増加し、社会問題になってきている。これらの病的状態の多くは、内臓脂肪の蓄積やインスリン抵抗性と同時に存在するため、これらの病気が一つにまとめられ、「シンドロームX (syndrome X)」、「死の四重奏」、「インスリン抵抗性症候群」、「内臓肥満型症候群」と呼ばれている。最近では、この症候群はメタボリックシンドローム(metabolic syndrome、MetS)と呼ばれるようになった。MetSに含まれる上記の病態は、アテローム性動脈硬化症や心臓血管疾患の重要な危険因子であり、これらの病態を同時に有していることは、心臓血管の疾患のリスクを付加的にではなく、相乗的に高めることが知られている。

30

【0008】

アディポネクチン(Adiponectin)は、抗糖尿病、抗動脈硬化、抗炎症性サイトカインと認められている。発明者および他のグループは、血清脂質濃度とインスリン抵抗性は血清アディポネクチン濃度と負の相関関係を示すことを報告している。また、発明者の報告を含むいくつかの報告により、血清高分子量アディポネクチン(high-molecular-weight (HMW) adiponectin)の測定は総アディポネクチン(total adiponectin)の測定より有用なことを示している。特に、薬物治療を受けている2型糖尿病患者では、アディポネクチンのHMW/totalの比率が単に血清総アディポネクチン濃度を測定するより有用であることが報告されている。

40

【課題を解決するための手段】

【0009】

50

本発明の発明者はMetS等の生活習慣病の発症予測にはインスリン抵抗性指数(HOMA-IR)が有用であること報告した。また、発明者は生活習慣病の発症とアディポネクチンの値の間に相関があることを見いだした。しかし、これらの項目は一般的な健康診断(健診)では扱われていない項目であり、これらの項目が生活習慣病の発症の予測にどのように関与するかを検証した事例は存在しない。

本発明の目的は、ANNシステムにより、幾つかの測定項目からなる臨床データから、何年後、例えば6年後にメタボリックシンドローム等の生活習慣病を発症するか否かの予測を可能とするものであり、従来のMLRによる分析より精度の高い予測を可能とする装置と方法及びプログラムを提供するものである。

【0010】

発明者は一般的な健診によって得られたデータに、アディポネクチンの値等、一般的な健診では測定されない項目を加えた診療データが生活習慣病の発症の予測にどのように関与するかを以下の方法により検証した。

- (1)：一般的な健診により得られる測定項目と、アディポネクチンの値等の他の項目から生活習慣病の発症に関与すると思われる18の項目を選択。
- (2)：健診者毎の診療データを収集し、学習用のデータと検証用のデータを生成。
- (3)：(1)で選択された全ての項目を含む学習用のデータにより、多重ロジスティック回帰(MLR)と人工ニューラルネットワーク(ANN)の学習を行う。
- (4)：学習が完了したMLRとANNに全ての項目を含む検証用のデータを入力し、生活習慣病を発症するか否かを予測する。
- (5)：(4)の予測結果と実際の診療データに基づいて判定した結果とを比較し、予測の感度(Sensitivity：陽性と判定されるべきものを正しく陽性と判定する確率)と特異度(Specificity：陰性と判定されるべきものを正しく陰性と判定する確率)を求める。
- (6)：1つあるいは複数の項目を除外した学習用のデータと検証用のデータを生成し、学習用のデータによりMLRとANNの学習を行う。
- (7)：学習が完了したMLRとANNに(6)で生成した検証用のデータを入力し、生活習慣病を発症するか否かを予測する。
- (8)：(7)の予測結果と実際の健診により得られたデータとを比較し、予測結果の感度と特異度を求める。
- (9)：(6)～(8)を全ての項目及び重要と思われる項目の組み合わせに対して実行する。
- (10)：(5)で得られた予測結果の感度と特異度と、(9)で得られた1つあるいは複数の項目を除外した時の予測結果の感度と特異度を比較し、各項目及び項目の組み合わせが予測結果の精度に寄与する度合いを検証する。

【0011】

上記の検証の結果以下のことが判った。

- (1)：生活習慣病の予測にはANNが優れている。
- (2)生活習慣病の発症予測には図3に示される18の項目(因子)全てが寄与するが、特に、体格指数(BMI)、年齢、拡張期血圧(DBP)、HDLコレステロール、LDLコレステロール、インスリン抵抗性指数(HOMA-IR)の6項目が重要である。
- (3)：メタボリックシンドロームの発症予測には、高分子量アディポネクチン、総アディポネクチン、インスリン、グリコアルブミン、および遊離脂肪酸(FFA)の各項目が関与している。
- (4)：耐糖能異常の発症予測には、血漿グルコースおよび(3)で挙げた6項目が関与している。
- (5)：脂質異常症の発症予測には、中性脂肪および(3)で挙げた6項目が関与している。

【図面の簡単な説明】

【0012】

【図1】本発明に用いたニューラルネットワークシステムの図

【図2】本発明の検討に用いた健診データの特性を示す図

【図3】本発明において用いた18の診療項目を示す図

10

20

30

40

50

- 【図4】健診データを多重ロジスティック回帰による分析した結果の図
- 【図5】学習用データセットと検証用のデータセットにおける健診者の特性を示す図
- 【図6】18の診療項目を用いた予測の感度と特異度を示す図
- 【図7】18の診療項目の感度を示す図
- 【図8-1】メタボリックシンドロームの発症予測の感度と特異度を示す図
- 【図8-2】メタボリックシンドロームの発症予測の感度と特異度を示す図
- 【図9】耐糖能異常の予測の感度と特異度を示す図
- 【図10】高血圧の発症予測の感度と特異度を示す図
- 【図11】脂質異常症の発症予測の感度と特異度を示す図
- 【図12】従来のニューラルネットワークシステムの図

10

【発明を実施するための形態】

【0013】

[ANNシステム]

ANNには、重層的なパーセプトロン(MLP)、放射状による関数ネットワーク(RBF)、および、リニアネットワーク(LIN)の3つのタイプのネットワークが知られている。MLPは図12に示されるように、1つの入力層と、1つ以上の中間層と、1つの出力層の各層から成り、出力層の中の1つのノードが結果を分類するために使用される。本実施例では100000個のMLPネットワークを生成した後、予測精度の高いネットワークを所定数選択して生活習慣病の発症確率の予測に用いた。

【0014】

20

図1は本発明において生活習慣病の発症確率の予測に用いたニューラルネットワークシステムの図である。ニューラルネットワークの構成は、入力層のノード数、中間層の層数、各中間層を構成するノード数、各重み係数の初期値、学習の方法(ノード間の伝搬関数)等の構成要素により決まり、構成要素の差異により多様な構成のニューラルネットワークを生成することが可能である。構成要素が異なるニューラルネットワークであっても、それらに同一の学習を行うことにより、多くの未知のパターンに対して同一の判定結果を出力するニューラルネットワークとすることが可能である。しかし、未知のパターンが「YES」と判定されるパターン領域と「NO」と判定されるパターン領域の境界の近傍に位置すると、構成要素の異なる複数のニューラルネットワークの一部が「YES」と判定し、他が「NO」と判定する場合がある。本発明で使用したニューラルネットワークシステムは、ニューラルネットワークの上記特性を利用するものであり、判定結果の確からしさ、例えば、疾患を発症する確率を、「YES」と判定したニューラルネットワークの割合により求めるものである。

30

図1の110-1ないし110-Mは、構成要素は異なるが、生活習慣病の発症に関して同一の学習を行ったM個のニューラルネットワークである。判定すべき診療データ(判定パターン)100をM個のニューラルネットワーク110-1ないし110-Mに並列に入力する。各ニューラルネットワークは生活習慣病の発症の可能性の有無を「YES」または「NO」により出力し、発症確率算出部111に与える。発症確率算出部111はM個の入力の内の「YES」の割合を算出し、生活習慣病が発症する確率112として出力する。

【0015】

40

[診断テスト]

本発明を検証するため、基準日に30~59歳であり、2000年と2006年の両方でその年の健康診断を受診した慶應義塾大学の410人の男性の診療データを評価した。なお、健診日に、MetS、内分泌疾患、重大な腎臓疾病あるいは肝臓疾病の罹患者、および糖尿病の薬物治療を受けている者は除外した。26人がスタチンを服用していたが、何れの被験者もニコチン酸を服用していなかった。

収縮期血圧(SBP)、拡張期血圧(DBP)および心拍数は、午前9時ごろ少なくとも5分間着席して休止した対象者から自動電子血圧計(BP-103III;日本コーリン、愛知県小牧市)を用いて2回測定した。

身長、体重、空腹時血糖(FPG)、血清中性脂肪(TG)、HDL-コレステロール、LDL-コレス

50

テロール、総コレステロール(TC)、グリコアルブミン予測インスリン、高分子量アディポネクチン(HMW-アディポネクチン)、および総アディポネクチン濃度は、夜間の絶食の後の午前9ごろ測定した。

体格指数(BMI)は体重(Kg)を身長(m)の2乗で除算したものである。血糖と血中脂質は通常の自動計測器によって分析した。血中インスリンは市販のキット(東ソー、東京)を用い、酵素免疫法にて測定した。日内変動と日差変動の係数は各々、2.9~4.6%と4.5~7.0%であった。インスリン抵抗性指数(HOMA-IR)はホメオスタシス・モデル評価(homeostasis model assessment)を用いて計算した。

血清総アディポネクチン濃度は、酵素免疫吸着法(ELISA法)(アディポネクチン ELISAキット、大塚製薬、東京)を使用して測定した。日内変動と日差変動の係数は<10%であった。また、高分子量アディポネクチンは、市販のキットを使用して測定した。HMW-アディポネクチンのためのELISAシステムは、たんぱく質を変性させる工程を必要とせず、また、使用されるモノクローナル抗体(IH7)は高分子量アディポネクチンに対して特異的に反応することが報告されている。

稀釈曲線は標準曲線と平行であった。この方法の日内変動と日差変動の係数は各々、2.4~3.0%と4.2~5.1%であった。

【0016】

[妥当性の確認]

バックプロパゲーションアルゴリズムによる学習プロセスを簡素化するために学習用のデータセットを使用し、アルゴリズムの進捗の検証のために検証用のデータセットを選択して使用した。

[統計分析]

統計分析は、Windows(商標)用のStatView(商標)(version5.0-J、SAS社、Cary、North Carolina)を使用して行った。P<0.05を統計学的に有意と考えた。血中中性脂肪、HOMA-IR、高分子量アディポネクチン、総アディポネクチン濃度は、それらの対数変換後に正規分布を示すため、これらのデータの対数を分析に使用した。2006年に健康診断を受けた慶應義塾大学の36~65歳の410人の男性受診者のウエスト周りとBMIの間の相関性が単純回帰分析によって決定された。

【0017】

[メタボリックシンドローム]

生活習慣病の予測に関する本発明を、6年後にメタボリックシンドローム(MetS)を発症するか否かの予測に適用した例について述べる。

BMI 24.3kg/m²に加えて、少なくとも次の3つの代謝異常の内の2つ以上が存在した時、MetSであるとした。なお、2000年の健診では臍周囲径(腹囲)の測定が行なわれていなかったため、日本人の診断の修正基準に基づいて変更して定義した。ウエスト周りと2006年のBMI(R²=0.709、p<0.0001、n=415)の相関図を使用すると、腹囲85cmは、BMIの24.3kg/m²に相当する。

(1)脂質異常：血清中性脂肪(TG) 150mg/dlおよび/またはHDLコレステロール<40mg/dl、または脂質低下薬による治療中のいずれか。

(2)高血圧：収縮期血圧 130mmHgおよび/または拡張期血圧 85mmHg、または降圧薬による治療中のいずれか。

(3)高血糖：空腹時血糖 110mg/dl、または抗糖尿病薬による治療中のいずれか。

【0018】

410人の日本人の男性(平均年齢は57.6年;年齢層は30-59年)の健康診断の結果が2000年から2006年まで追跡され(図2)、410人のうち27人(6.6%)が6年後にMetSと診断された。

【0019】

[入力項目と出力]

図3はMLRとANN分析を使用して各健診者の6年後の状況を予測するため使用した入力項目と出力を示すものであり、入力項目のX₁~X₈として以下の診療データを使用した。

X¹: 年齢(歳)

10

20

30

40

50

- X_2 : 体格指数 (BMI : Kg/m²)
 X_3 : 収縮期血圧 (SBP : mmHg)
 X_4 : 拡張期血圧 (DBP : mmHg)
 X_5 : 血漿グルコース (mg/dl)
 X_6 : 中性脂肪 (mg/dL)
 X_7 : 高密度リポ蛋白 - コレステロール (mg/dL)
 X_8 : 低密度リポ蛋白 - コレステロール (mg/dL)
 X_9 : 血清AST (IU/L)
 X_{10} : 血清ALT (IU/L)
 X_{11} : 高分子量 (HMW) アディポネクチン (μg /mL)
 X_{12} : 総 (total) アディポネクチン (μg /mL)
 X_{13} : グリコアルブミン (%)
 X_{14} : 総コレステロール (mg/dL)
 X_{15} : 遊離脂肪酸 (mEq/L)
 X_{16} : インスリン (μU /mL)
 X_{17} : インスリン抵抗指数 (HOMA-IR)
 X_{18} : 喫煙 (0 = なし、1 = あり)

10

【 0 0 2 0 】

X_1 と X_2 はそれぞれ健診者の年齢と体格指数 (BMI) を表している。 X_3 と X_4 はそれぞれ収縮期血圧と拡張期血圧を表している。 $X_5 - X_{17}$ は検査所見を表し、 X_5 は血漿グルコース、 X_6 は中性脂肪、 X_7 はHDLコレステロール、 X_8 はLDLコレステロール、 X_9 は血清AST、 X_{10} は血清ALT、 X_{11} は高分子量 (HMW) アディポネクチン、 X_{12} は総コレステロール、 X_{13} はグリコアルブミン、 X_{14} は総アディポネクチン、 X_{15} は遊離脂肪酸、 X_{16} はインスリン、 X_{17} はHOMA-IRである。また、 X_{18} は喫煙の有無を表している。

20

出力は2006年にMetSを発症するが否かである。

【 0 0 2 1 】

図4は多重ロジスティック回帰 (MLR) による分析の結果である。同図からメタボリックシンドロームの発生確率の予測には、項目 X_2 (BMI) が最も高い有意性 ($p=0.0002$) を有していることがわかる。また、項目 X_7 (HDL-C) と X_4 (DBP) は有効な項目であるが、統計的には有意ではなかった (それぞれ $p=0.0505$ と 0.0648)。他の項目は、MetSの発症にはほとんど影響しないとする結果になっている。

30

【 0 0 2 2 】

前述の16の項目を用いて、MLRとANNにより予測を行った。予測に当たり、全体の診療データを学習用のデータセット ($n=246$) と検証用のデータセット ($n=164$) にランダムに分割した。図5に示されるように、すべての項目において学習用のデータセットと検証用のデータセットの間の有意な差は存在しなかった。

【 0 0 2 3 】

図6(a)はMLRとANNにより得られた6年後の発症予測と、6年後の診療データに基づいて判定された発症 / 非発症の結果との差を感度と特異度により示すものである。同図に示されるように、MLRとANNの特異度は各々0.95と0.91であり、大きな差は認められない。しかし、MLRの感度は0.27であるのに対してANNの感度は0.93であり、ANNの精度が高い。

40

ANNでは感度と特異度の値が類似している。この結果は、ANNを使用することにより、利用価値のある予測を行うことが可能であることを示している。

訓練データと確認データの組み合わせを変更して更に4回の検証を行った。図6(b)に各検証の感度と特異度の数値が示されているが、全ての場合においてANNの感度はMLRの感度より著しく大きかった ($p<0.005$)。

【 0 0 2 4 】

ANNは有用なモデルであるが、非線形の方法でデータを処理するため、推論のネットワーク・ロジックを単純な要素に分割することができない問題がある。この問題を扱うために感度分析 (SA) を使用して項目の重要度を調べた。

50

SAはテスト項目(X_{test})を除いた場合に予測結果の確度がどのように変わるかを測定するものである。値が「1」以上の X_{test} は予測を改善することを示し、値が「1」未満の項目は予測結果を改善しないことを示している。全てのニューラルネットワークを分析して平均値と標準偏差を確認した。

【0025】

図7に感度分析(SA)の結果を示す。分析結果からBMI(X_2)が最も重要な項目であること認識される。また、MLR分析(図4)では重要な関係を持っていなかった、年齢(X_1)、DBP(X_4)、HDLコレステロール(X_7)、LDLコレステロール(X_8)、HOMA-IR(X_{17})の5つの項目が高い感度を示している。これはBMI以外の要素は6年後のMetSの発症と非線形の関係があることを示している。上記結果から、年齢(X_1)、BMI(X_2)、DBP(X_4)、HDLコレステロール(X_7)、LDLコレステロール(X_8)、HOMA-IR(X_{17})の6つの項目がMetSの発症予測に重要であることが見いだされた。

10

【0026】

図8-1は、図3に示される18項目の全てを用いてANNの学習と検証を行って得られた感度と特異度と、前記の6項目を除外したデータを用いてANNの学習と検証を行って得られた感度と特異度を示すものである。

(1)に示される様に、全項目を用いた場合の感度と特異度は0.96と0.91であるのに対して、(2)に示される6項目を除外した場合の感度と特異度は0.88と0.84に低下している。当該結果はMetSの予測に、年齢、BMI、DBP、HDLコレステロール、LDLコレステロール、HOMA-IRの6項目が重要であることを示している。

20

【0027】

次にメタボリックシンドロームの発症予測に寄与していると思われる他の項目を除外したデータによりANNの学習と検証を行い、除去した項目の感度と特異度を検証した。図8-2は、HOMA-IR、高分子量アディポネクチン、総アディポネクチン、インスリン、グリコアルブミン、および遊離脂肪酸(FFA)の第2の項目のグループを除外した場合の感度と特異度である。当該項目を除去した感度は0.78であり、図8-1に示される、全ての項目を含むデータにより検証した感度0.96より大きく低下している。この結果から、MetSの発症の予測に当該項目が寄与していることが判る。

【0028】

メタボリックシンドローム(MetS)を発症する危険因子に関して多くの報告が存在する。しかし、疫学的な研究および医学的統計に由来したデータは、必ずしも個人レベルでMetSの発症を正確に予測することが可能とは言えない。上記の検証の結果から、ANNは個人の予測に良い結果を出していることが判る。図4に示されるように、MLRは特徴的な項目を見つけ、BMIや血清HDLコレステロールのような有意な要素を特定することに有用である。しかし、MLRは感度が非常に低く、本発明で採用した項目を使用する予測には適切ではなかった。原因と結果の間に非線形の相関性が存在する場合、従来のMLRに対してANNは理論的な長所を有していることが見いだされた。

30

【0029】

[耐糖能異常]

図9はANNを用いて生活習慣病の一つである耐糖能異常の発症を予測した場合の感度と特異度を示すものである。先のMetSの学習と検証に用いた各診療データの被験者が、6年後に耐糖能異常と診断されたか否かを示すデータを正解データとし、ANNの学習と検証を行った。

40

図9の(1)に示される様に、図3に示される18の全ての項目を用いて学習と検証を行った場合の感度と特異度は、0.75と0.85である。また、図9の(2)に示される様に、年齢、BMI、DBP、HDLコレステロール、LDLコレステロール、HOMA-IRの6項目を除外して学習と検証を行った場合の感度と特異度は0.73と0.80である。当該結果から前記6項目が耐糖能異常の発症の予測に寄与していることが判る。

図9の(3)は、前記6項目と血漿グルコース($Glc : X_5$)を除外して学習と検証を行った場合の感度と特異度である。当該検証における感度と特異度は0.52と0.65であり、血漿グ

50

ルコース(Glc)が耐糖能異常の発症の予測に寄与していることが判る。

【0030】

[高血圧]

図10はANNを用いて生活習慣病の一つである高血圧の発症を予測した場合の感度と特異度を示すものである。先のMetSの学習と検証に用いた各診療データの被験者が、6年後に高血圧と診断されたか否かを示すデータを正解データとし、ANNの学習と検証を行った。

図10の(1)に示される様に、図3に示される18項目の全てを用いて学習と検証を行った場合の感度と特異度は0.84と0.76である。また、図10の(2)に示される様に、年齢、BMI、DBP、HDLコレステロール、LDLコレステロール、HOMA-IRの6項目を除外して学習と検証を行った場合の感度と特異度は0.66と0.63である。当該データから前記6項目が高血圧の発症の予測に寄与していることが判る。

【0031】

[脂質異常症]

図11はANNを用いて生活習慣病の一つである脂質異常症の発症を予測した場合の感度と特異度を示すものである。先のMetSの学習と検証に用いた各診療データの被験者が、6年後に脂質異常症と診断されたか否かを示すデータを正解データとし、ANNの学習と検証を行った。

図11の(1)に示される様に、図3に示される18項目の全てを用いて学習と検証を行った場合の感度と特異度は0.68と0.76である。また、図11の(2)に示される様に、年齢、BMI、DBP、HDLコレステロール、LDLコレステロール、HOMA-IRの6項目を除外して学習と検証を行った場合の感度と特異度は0.69と0.70である。当該データから前記6項目が脂質異常症の発症の予測、特に特異度の値に影響していることが判る。

図11の(3)は、前記6項目と中性脂肪(TG: X_6)を除外して学習と検証を行った場合の感度と特異度である。当該検証における感度と特異度は0.62と0.66であり、中性脂肪(TG)が脂質異常症の発症の予測に寄与していることが判る。

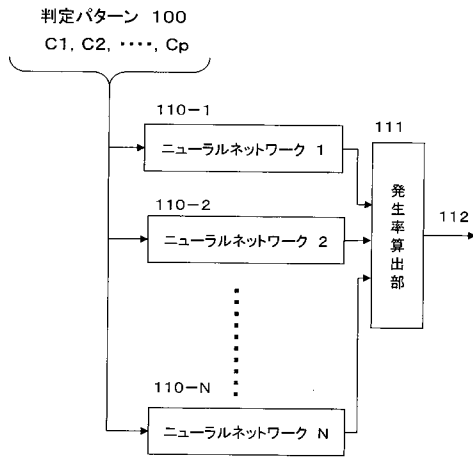
【産業上の利用可能性】

【0032】

本発明の検証結果はANNが健診者個々人の生活習慣病の発症の可能性を予測する有効な方法であることを示している。健康診断により健康の維持に問題が在ると診断された健診者に対して、生活習慣を変更し、それを維持するように動機づけることは難しい。しかし、本発明を用いることにより健康診断の結果から生活習慣病の危険性を予測し、医師が各健診者にライフスタイル等に関して助言する際の手助けとすることが可能となる。また、健診者個人にMetS等の発症の可能性を示すことにより、生活習慣を変える高い動機づけとなり、生活習慣の変更の契機となり得る。さらに、生活習慣を変更した場合のデータを入力し、その結果を予測することにより、MetS等の発症の可能性がどのように変化するかをシミュレートすることが可能であり、健診者の関心を強くすることが可能となる。

また、MetS等の生活習慣病の発生の可能性は未知の因子に影響されている可能性もある。しかし、データを更新しANNの学習を継続的に行うことにより、それらの問題に対応可能であり、これにより医療のコストを縮小することが可能となる。

【 図 1 】



【 図 2 】

健診者の特性	
Total n=415	
Age (years)	44.2 (30-59)
BMI (kg/m ²)	22.7 (17.3-32.2)
SBP (mmHg)	119.4 (90-170)
DBP (mmHg)	75.4 (51-111)
Glucose (mg/dL)	92.2 (60-146)
TG (mg/dl)	110.5 (18-424)
HDL-cholesterol (mg/dL)	55.1 (27-118)
LDL-cholesterol (mg/dL)	123.5 (32-213)
AST (IU/L)	23.7 (12-75)
ALT (IU/L)	25.8 (6-131)
HMW-adiponectin (μg/mL)	4.7 (0.3-17.86)
Total adiponectin (μg/mL)	6.6 (0.6-19.6)
Glycated albumin (%)	16.4 (12.5-35.1)
Total cholesterol (mg/dL)	194.8 (97-295)
FFA (mEq/L)	0.64 (0.18-1.60)
Insulin (μU/mL)	4.9 (0-22)
HOMA-IR	1.13 (0.5-43)
Smoking habit (number)	No (370), Yes (45)
MetS in 2006 (number)	No (388), Yes (27)

連続したデータは平均値が示され、括弧内に範囲または個数が記載されている

BMI: 体格指数、SBP: 収縮期血圧、DBP: 拡張期血圧、TG: トリグリセリド、HDL: 高密度リポタンパク質、LDL: 低密度リポタンパク質、AST: アスパラギン酸アミノ基転移酵素、ALT: アラントランスアミナーゼ、HMW: 高分子量、GA: グリコアルブミン、FFA: 遊離脂肪酸、HOMA-IR: ホメオスタシス-モデル評価インスリン抵抗指数

【 図 3 】

予測に用いた項目(項目)と出力	
Factor (項目)	
X ₁ : Age (years)	X ₁₀ : ALT (IU/L)
X ₂ : Body mass index (kg/m ²)	X ₁₁ : HMW-adiponectin (μg/mL)
X ₃ : SBP (mmHg)	X ₁₂ : Total adiponectin (μg/mL)
X ₄ : DBP (mmHg)	X ₁₃ : Glycated albumin (%)
X ₅ : Blood glucose (mg/dl)	X ₁₄ : Serum total cholesterol (mg/dL)
X ₆ : Serum triglycerides (mg/dL)	X ₁₅ : Free fatty acid (mEq/L)
X ₇ : HDL-cholesterol (mg/dL)	X ₁₆ : Insulin (μU/mL)
X ₈ : LDL-cholesterol (mg/dL)	X ₁₇ : HOMA-IR
X ₉ : AST (IU/L)	X ₁₈ : Smoking: 0=no, 1=yes
Outcome (出力)	

6年後の代謝症候群の発症

0=no

1=yes

【 図 4 】

多重ロジスティック回帰による分析結果				
Factor (項目)	Regression coefficient	Standard error	χ ² value	p value
Constant	35.8560	9.3511	14.7027	0.0001
X ₁	-0.0528	0.0393	1.8085	0.1787
X ₂	-0.6092	0.1615	14.2281	0.0002
X ₃	-0.0255	0.0304	0.7011	0.4024
X ₄	-0.0770	0.0417	3.4094	0.0648
X ₅	-0.0874	0.0680	1.6533	0.1985
X ₆	-0.0092	0.0074	1.5786	0.2090
X ₇	0.1046	0.0535	3.8249	0.0505
X ₈	-0.0223	0.0394	0.3205	0.5713
X ₉	0.0241	0.0713	0.1142	0.7354
X ₁₀	0.0401	0.0338	1.4079	0.2354
X ₁₁	0.2056	0.3347	0.3774	0.5390
X ₁₂	0.0082	0.1847	0.0020	0.9644
X ₁₃	-0.1289	0.1356	0.9031	0.3420
X ₁₄	0.0137	0.0397	0.1196	0.7295
X ₁₅	-1.0277	1.2263	0.7023	0.4020
X ₁₆	-1.3309	1.0102	1.7357	0.1877
X ₁₇	5.1979	4.2995	1.4615	0.2267
X ₁₈ [0]	-1.1295	0.7975	2.0058	0.1567

【 図 5 】

学習用データセットと検証用のデータセットにおける健診者の特性

	Training (n 246) (学習用)	Validation (n 164) (検証用)
Age (years)	44.2 (30-59)	44.2 (30-59)
BMI (kg/m ²)	22.8 (17.6-30.6)	22.7 (17.3-32.2)
SBP (mmHg)	119.0 (90-170)	120.0 (93-166)
DBP (mmHg)	74.8 (51-111)	76.3 (52-111)
Glucose (mg/dL)	92.1 (77-146)	92.2 (60-127)
TG (mg/dL)	110.1 (29-424)	111.1 (18-395)
HDL-cholesterol (mg/dL)	55.2 (27-118)	54.9 (32-104)
LDL-cholesterol (mg/dL)	121.7 (51-195)	126.2 (32-213)
AST (IU/L)	23.5 (12-57)	24.1 (13-75)
ALT (IU/L)	25.4 (6-118)	26.5 (7-131)
HMW-adiponectin (µg/mL)	4.7 (1.11-17.86)	4.6 (0.25-14.26)
Total adiponectin (µg/mL)	6.6 (1.63-19.6)	6.6 (0.63-17.8)
Glycated albumin (%)	16.4 (12.5-35.1)	16.4 (12.5-22.0)
Total cholesterol (mg/dL)	194.8 (120-287)	199.0 (97-295)
FFA (mEq/L)	0.64 (0.18-1.33)	0.65 (0.21-1.60)
Insulin (µU/mL)	4.9 (0-22)	5.0 (0-22)
HOMA-IR	1.12 (0-5.07)	1.15 (0-5.43)
Smoking habit (number)	No (223), Yes (23)	No (142), Yes (22)
MetS in 2006 (number)	No (230), Yes (16)	No (153), Yes (11)

数値は平均値(範囲)、或いは項目の数
学習用のデータと検証用のデータの間には特筆すべき差異は存在しない

【 図 6 】

多重ロジスティック回帰(MLR)とニューラルネットワーク(ANN)の感度(Sensitivity)と特異度(Specificity)

	MLR	ANN
感度 (Sensitivity)	0.27	0.93
特異度 (Specificity)	0.95	0.91

(a)

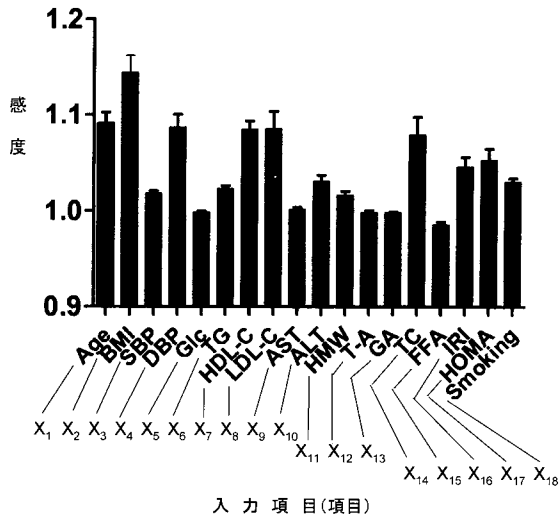
多重ロジスティック回帰(MLR)とニューラルネットワーク(ANN) 感度と特異度の他の測定結果

MLR	Trial 1	Trial 2	Trial 3	Trial 4
感度 (Sensitivity)	0.40	0.50	0.25	0.29
特異度 (Specificity)	0.95	0.93	0.99	0.97

ANN	Trial 1	Trial 2	Trial 3	Trial 4
感度 (Sensitivity)	0.93	0.93	0.85	0.96
特異度 (Specificity)	0.90	0.87	0.90	0.88

(b)

【 図 7 】



【 図 8 - 1 】

	(1)全項目	(2)6項目を除く
感度	0.96	0.88
特異度	0.91	0.84

メタボリックシンドロームの発症予測の確度

【 図 8 - 2 】

感度	0.78
特異度	0.95

第2グループの項目をのぞいたデータから得られたデータニューラルネットワーク(ANN)による感度と特異度

【 図 9 】

	(1)全項目	(2)6項目を除外	(3)6項目とGlcを除外
感度	0.75	0.73	0.52
特異度	0.83	0.80	0.65

耐糖能異常の発症予測の確度

【 図 1 1 】

	(1)全項目	(2)6項目を除外	(3)6項目とTGを除外
感度	0.68	0.69	0.62
特異度	0.76	0.70	0.66

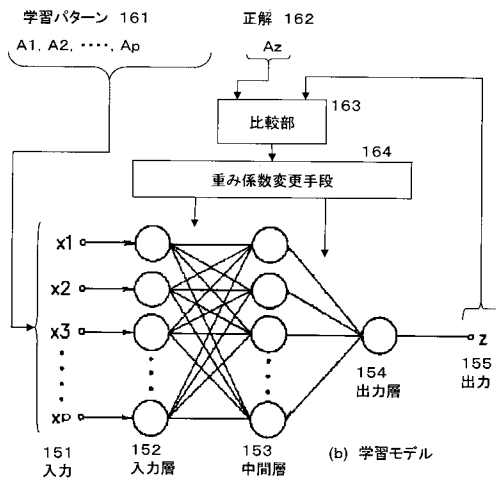
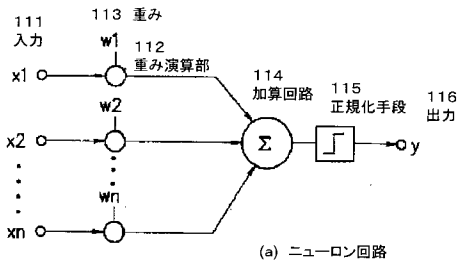
脂質異常症の発症予測の確度

【 図 1 0 】

	(1)全項目	(2)6項目を除外
感度	0.84	0.66
特異度	0.76	0.63

高血圧の発症予測の確度

【 図 1 2 】



フロントページの続き

(72)発明者 朴沢 重成

東京都新宿区信濃町3番地 慶應義塾大学 医学部内

(72)発明者 広瀬 寛

神奈川県横浜市港北区日吉4丁目1番1号 慶應義塾大学 日吉キャンパス内

(72)発明者 高山 哲朗

東京都新宿区信濃町3番地 慶應義塾大学 医学部内

Fターム(参考) 4C117 XA05 XB02 XB09 XB12 XE54 XJ34 XJ38 XR01

专利名称(译)	生活方式相关疾病的诊断预测装置，生活方式疾病的诊断预测方法和程序		
公开(公告)号	JP2012064087A	公开(公告)日	2012-03-29
申请号	JP2010208988	申请日	2010-09-17
[标]申请(专利权)人(译)	学校法人庆应义塾		
申请(专利权)人(译)	学校法人庆应义塾		
[标]发明人	朴沢重成 広瀬寛 高山哲朗		
发明人	朴沢 重成 広瀬 寛 高山 哲朗		
IPC分类号	G06Q50/22 A61B5/00 G06N3/00		
FI分类号	G06F17/60.126.G A61B5/00.G G06N3/00.550.Z G06N20/00 G06N3/04.127 G06N99/00.150 G06Q50/22 G06Q50/22.106 G16H20/00 G16H50/00		
F-TERM分类号	4C117/XA05 4C117/XB02 4C117/XB09 4C117/XB12 4C117/XE54 4C117/XJ34 4C117/XJ38 4C117/XR01 5L099/AA04		
代理人(译)	高桥秀夫		
外部链接	Espacenet		

摘要(译)

要解决的问题：使用人工神经网络预测几年后是否有可能发展与生活方式有关的疾病。包括具有输入层，至少一个中间层和输出层的三层网络的神经网络至少用于年龄，体重指数（BMI），舒张压（DBP），HDL胆固醇，LDL胆固醇和HOMA-IR。通过学习由6个项目组成的学习模式来学习发展与生活方式有关的疾病的可能性，从健康检查等获得的医学数据，几年后是否有可能发展与生活方式有关的疾病的可能性预测。

[选型图]图1

