

# 個人のエネルギー消費・摂取量と体脂肪率の相関を特徴付けるパラメータに関する研究

竹内 裕之<sup>†</sup> 黛 勇氣<sup>‡</sup> 児玉 直樹<sup>†</sup>

<sup>†</sup>高崎健康福祉大学 健康福祉学部 医療情報学科 〒370-0033 群馬県高崎市中大類町 37-1

<sup>‡</sup>高崎健康福祉大学大学院 健康福祉学研究科 〒370-0033 群馬県高崎市中大類町 37-1

E-mail: <sup>†</sup>(htakeuchi, kodama)@takasaki-u.ac.jp, <sup>‡</sup>0910404@takasaki-u.ac.jp

**あらまし** 個人の日々のエネルギー消費・摂取量と体脂肪率変化の時系列相関解析において、最大の相関係数をもたらすエネルギー消費・摂取量の時系列加算処理の重み付パターンを求めた。重み付パターンは2つのパラメータで表現され、その組み合わせが個人の健康に関する特質を表すとの想定のもとに解析を行った結果、(1)ある個人の長期的な時系列データにおいて、加齢や季節に関係なくパラメータの組み合わせはある偏りを示す、(2)複数の被験者の時系列データでは、パラメータの組み合わせは大きく2つのグループに分かれる傾向がある、ことなどが判った。

**キーワード** 時系列相関解析, エネルギー消費・摂取量, 体脂肪率変化、重み付パターン

## Study on Parameters Characterizing the Nature of Personal Health in the Correlation between Energy Expenditure/Supply and Body-fat Percentage

Hiroshi TAKEUCHI<sup>†</sup> Yuuki MAYUZUMI<sup>‡</sup> and Naoki KODAMA<sup>†</sup>

<sup>†</sup> Department of Healthcare Informatics, Faculty of Health and Welfare, Takasaki University of Health and Welfare  
37-1, Nakaorui-machi, Takasaki-shi, Gunma, 370-0033 Japan

<sup>‡</sup> Post-Graduate Course of Health and Welfare, Graduate School of Takasaki University of Health and Welfare, 37-1,  
Nakaorui-machi, Takasaki-shi, Gunma, 370-0033 Japan

E-mail: <sup>†</sup>(htakeuchi, kodama)@takasaki-u.ac.jp, <sup>‡</sup>0910404@takasaki-u.ac.jp

**Abstract** The weighting patterns for the summation of daily time-series energy expenditure and supply data giving the maximum correlations to the variation of daily body-fat percentage data were obtained. The weighting patterns can be expressed by two parameters whose combination is supposed to characterize the nature of personal health. The combination of the parameters for a subject was found to show a significant bias in the frequency distribution, independent of season and aging, for the term of seven years, and the combination of the parameters of 20 other subjects showed a tendency to divide into two types.

**Keyword** Time-series Correlation Analysis, Energy Expenditure/Supply, Variation of Body-fat Percentage, Weighting Pattern

### 1. はじめに

少子高齢化社会の我が国において、自身の健康管理を徹底して生活習慣病を予防することは喫緊の課題である。近年この分野での情報技術の活用が期待されている[1,2]。我々は、携帯電話(mobile phone)を情報端末として、個人の日常の生活習慣や健康に関するデータをインターネット上のサーバに蓄積し、サーバ上で生活習慣と健康状態の相関ルール抽出(健康データマイニング)を行って個人に通知する、個人健康管理システムを開発してきた[3]。健康管理を的確に行うには、

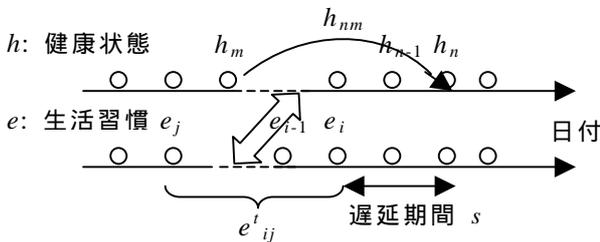
生活習慣と健康状態の関わりにおいて個人の特質を把握することが肝要と考えられる。開発した健康データマイニングのアルゴリズムでは、個人毎に蓄積された生活習慣と健康状態に関わる時系列データ解析が中核になっているが、先の研究で、生活習慣データの時系列加算処理に重み付けの概念を導入した[4]。

本報では、健康状態の変化と生活習慣データの時系列加算の間に最大の相関をもたらす重み付けのパターンが個人の特質を反映しているという想定のもとに行ったデータ解析の結果を報告する。

## 2. 研究方法

### 2.1. 時系列データ解析

我々が開発した健康データマイニングでは、「生活習慣の蓄積が健康状態に変化をもたらし、その影響は時間遅れをもって現れることがある」という極めてシンプルなモデルをベースとしている[5]。すなわち、ある健康状態の変化を目的変数とし、時間遅れを考慮したある期間の生活習慣の蓄積を入力変数として相関ルールを抽出する。相関ルールの抽出には、あらかじめ、ある期間の時系列データを基にして、目的変数である健康状態の変化に影響を及ぼす生活習慣の蓄積をスクリーニングする。スクリーニングには図1に示す時系列データ間のピアソンの積率相関係数を用いる。



$$\text{相関係数} : r(\Delta h_{nm}, e^{t_{ij}}) = \frac{\text{Cov}(\Delta h_{nm}, e^{t_{ij}})}{SD(\Delta h_{nm})SD(e^{t_{ij}})}$$

$$\text{健康状態の変化} : \Delta h_{nm} = h_n - h_m$$

$$\text{生活習慣の蓄積} : e^{t_{ij}} = e_i + e_{i-1} + \dots + e_j$$

図1. 生活習慣と健康状態の時系列相関の評価

図1において、 $SD(h_{nm})$ は健康状態の変化  $h_{nm}$ の時系列区間における標準偏差、 $SD(e^{t_{ij}})$ は生活習慣の蓄積  $e^{t_{ij}}$ の時系列区間における標準偏差、 $\text{Cov}(h_{nm}, e^{t_{ij}})$ は  $h_{nm}$ と  $e^{t_{ij}}$ の共分散である。時間遅れは遅延期間  $s = n - i \geq 1$ で表現する。

実際のスクリーニングは、対象とする健康状態  $h$ と各種生活習慣  $e$ の時系列データについて、 $n-m, i-j, s$ をパラメータとして相関を評価する。すなわち、各  $(n-m, i-j, s)$ のセットにつき相関係数を計算し、もし、1つ以上の相関係数の絶対値がある閾値  $r_s$ より大きければ、その  $e$ の蓄積を  $h$ に対する入力変数として採用する。そして、データマイニング処理における実際の入力フィールドは相関係数の絶対値が最大となる  $(n-m, i-j, s)$ のセット  $((n-m)_{\max}, (i-j)_{\max}, s_{\max})$ をもとに定義する。例えば、 $(i-j)_{\max}=2, s_{\max}=2$ であれば、 $e$ に関わる入力フィールドを

$$e_i + e_{i-1} + e_{i-2} \quad (i=n-2)$$

と定義する。すなわち、「2日前から3日間の生活習

慣  $e$ の蓄積」を入力変数のひとつとして定義する。ここで、 $(i-j)_{\max}$ が大きいということは、長期間の生活習慣の蓄積が現在の健康状態に影響を与え、 $s_{\max}$ が大きいということは、生活習慣の蓄積が遅れをもって現在の健康状態に影響を与えるということになる。

### 2.2. 加算処理への重み付けの導入

これまでの時系列データ解析において、何日間かの生活習慣データを単純加算し、加算日数と遅延期間をパラメータとして時系列相関を評価してきた(図1)。しかし、生活習慣データを過去に遡って日毎に単純に加算するというのは、健康状態との関わりにおいて複雑な生体反応をあまりにも単純化しすぎていると考えられる。むしろ、過去の生活習慣が現在の健康状態に変化を与えるという観点からは、即効性のあるもの遅効性のあるものなど様々なパターンの影響が考えられる。実際に、これまでの研究から遅延期間を考慮して加算すると相関係数の絶対値が著しく増大するケースが多々みられた[6]。

そこで、過去の生活習慣データを適当な重みをつけて加算することで、より精度の高い時系列相関解析ができるのではないかと期待が出てくる。即ち、適当な重み付けをすることにより、より高い相関係数をもつ条件を見つけ出すことが出来ないかという期待である。先の研究で、これまでの時系列データ解析結果を参考に、図2に示した正規分布関数に着目し、 $\mu$ と  $\sigma$ の2つのパラメータで重み付けを特徴付ける方法を考案した[4]。従来の単純加算方法と比較すると、加算日数が大きいことは  $\mu$ が大きいことに対応し、遅延日数が大きいことは  $\sigma$ が大きいことに対応する。ここで、正規分布関数を重み付け関数として使うときは、 $x$ の値はゼロもしくは正の整数値として得られた関数値を重みとして過去の日付に降順に割り当てる(図2の重み付け加算の例を参照)。

#### [重み付け関数の定義]

$$w(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (\sigma > 0, \mu \geq 0)$$

$$x = 0, 1, 2, \dots, m$$

$$\sum_{i=0}^m w(i) = 1$$

#### [生活習慣データ重み付け加算の例]

$$w(0)e_{n-1} + w(1)e_{n-2} + w(2)e_{n-3} + \dots + w(10)e_{n-11}$$

( $m = 10$ , 健康データ日付のカーソル:  $n$ と仮定)

図2. 採用した重み付け関数

いくつかの  $\mu$  と  $x$  の値に対応する重み付け関数のパターンを図3に示す。ここでは、健康データを取得した日の前日から遡って11日間の生活習慣データに重みを付けて加算することを想定している。

時系列相関評価の結果、図3(a)のようなパターンの重み付けで相関係数の絶対値が最大になれば、比較的直近の生活習慣が遅延無く健康状態に影響を与えることになり、(b)や(d)のようなパターンの重み付けで相関係数の絶対値が最大になれば、比較的長期の生活習慣が尾を引いていることになる。また、(c)のようなパターンの重み付けで相関係数の絶対値が最大になれば、比較的短期間の生活習慣が遅れをもって健康状態に影響を与えていることになる[4]。

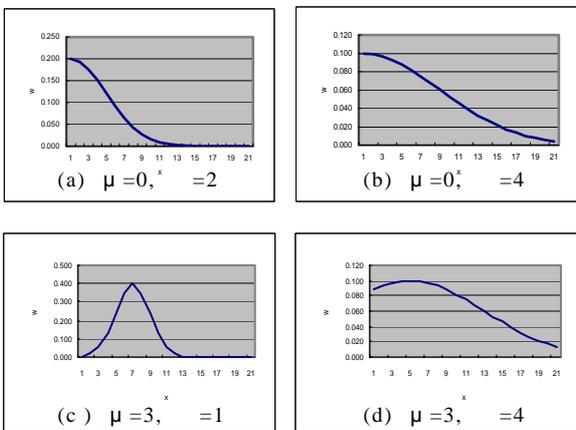


図3. 各種  $\mu$  と  $x$  の値に対応する重み付け関数のパターン

### 2.3. 対象データと取得方法

本研究では、まず生活習慣との関わりにおいて個人の健康に関する特質が加齢や季節による影響を受けるかどうかを調べるために、データ取得開始時58歳男性（通いで通常勤務）がその後7年間余りに亘りほぼ毎日取得した体脂肪率、消費カロリーおよび摂取カロリーの時系列データを対象に解析した。さらに、個人差が現れるかどうかを調べるために20人のボランティア（平均年齢22歳、男性9名、女性11名）の半年に亘る体脂肪率と消費カロリーの時系列データを対象に解析した。

ここで、体脂肪率は毎朝体組成計(タニタ製)を用いて取得され、消費カロリーは歩数計(オムロン製)で計測された日毎の歩数消費カロリーとその他運動による消費カロリーの和である。また、摂取カロリーは、朝、昼、夜、それぞれの食事内容から換算した摂取エネルギーの日毎の総和である。

## 3. 解析結果

### 3.1. 1被験者の7年間に亘るデータ解析

#### 3.1.1. 体脂肪率と消費カロリーの相関

解析に用いた重み付け関数は、図2において、 $\mu=0, 1, 2, 3, \dots, x=0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0$  の各組み合わせの32種類のパターンである。ここで、 $\mu$  の値を整数値としたのは、単純加算による従来法における遅延日数とほぼ等価なパラメータであるがためである。一方、 $x$  は従来法における加算日数そのものでなく、その大小の尺度パラメータに相当するので小刻みに設定した。 $x$  の範囲は  $x=0\sim 10$  で、体脂肪率を取得した日の前日から11日間の総消費カロリーデータの重み付け加算を行った。

対象データ取得期間は2004年6月1日～2011年8月31日の7年3か月間で、3カ月毎(季節毎)に時系列相関係数を評価し、相関係数の絶対値が最大になる( $\mu, x$ )の組み合わせを求めた。得られた結果を相関係数の値  $r$  とともに表1にまとめた。最大の相関係数が0.20未満のデータ期間では( $\mu, x$ )の値は null と表記している。

表1. 消費カロリーと体脂肪率の時系列データ間の相関係数が最大となる( $\mu, x$ )のセット

データ期間	$\mu$	$r$	データ期間	$\mu$	$r$	データ期間	$\mu$	$r$
04/6/1-8/31	(1, 0.5)	-0.403	05/6/1-8/31	(1, 0.5)	-0.253	06/6/1-8/31	(3, 2.5)	-0.343
04/9/1-11/30	(3, 1)	-0.341	05/9/1-11/30	(3, 1.5)	-0.463	06/9/1-11/30	(1, 0.5)	-0.216
04/12/1-05/2/28	(2, 1)	-0.329	05/12/1-06/2/28	(2, 4)	-0.462	06/12/1-07/2/28	(1, 0.5)	-0.315
05/3/1-5/31	(2, 1)	-0.482	06/3/1-5/31	(2, 0.5)	-0.272	07/3/1-5/31	(3, 4)	-0.577
データ期間	$\mu$	$r$	データ期間	$\mu$	$r$	データ期間	$\mu$	$r$
07/6/1-8/31	(2, 1.5)	-0.474	08/6/1-8/31	(3, 3.5)	-0.369	09/6/1-8/31	(3, 1.5)	-0.346
07/9/1-11/30	(0, 1)	-0.2	08/9/1-11/30	(2, 1)	-0.479	09/9/1-11/30	(3, 1.5)	-0.482
07/12/1-08/2/28	(1, 0.5)	-0.2	08/12/1-09/2/28	(2, 0.5)	-0.359	09/12/1-10/2/28	null	
08/3/1-5/31	(1, 1)	-0.509	09/3/1-5/31	(1, 1)	-0.336	10/3/1-5/31	(1, 0.5)	-0.291
データ期間	$\mu$	$r$	データ期間	$\mu$	$r$			
10/6/1-8/31	(1, 0.5)	-0.261	11/6/1-8/31	(2, 1.5)	-0.42			
10/9/1-11/30	(2, 0.5)	-0.31						
10/12/1-11/2/28	null							
11/3/1-5/31	(2, 0.5)	-0.349						

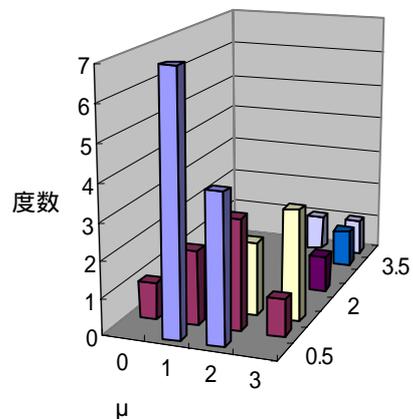


図4. 消費カロリーと体脂肪率の時系列データ間の相関係数が最大となる( $\mu, x$ )のセットの度数分布(被験者1名、7年間計27セット)

当初、最大の相関を示す ( $\mu$ ,  $r$ ) について加齢による変化や季節による変動などを予想していたが、これらに関係なくばらついており、 $\mu = 1, 2, r = 0.5, 1.0$  の組み合わせの頻度が高く、割合が 16/27 (60%) となっている (図 4)。

### 3.1.2. 体脂肪率と摂取カロリーの相関

体脂肪率と消費カロリーとの相関解析と全く同じ条件で、解析を行った結果を表 2 にまとめた。

最大の相関を示す ( $\mu$ ,  $r$ ) の組み合わせは、やはり加齢や季節とは関係なくばらついており、 $\mu = 2, 3, r = 0.5, 1.0$  の組み合わせの頻度が高く、割合が 16/22 (73%) となっている (図 5)。

表 2 . 摂取カロリーと体脂肪率の時系列データ間の相関係数が最大となる ( $\mu$ ,  $r$ ) のセット

データ期間	( $\mu$ , $r$ )	データ期間	( $\mu$ , $r$ )	データ期間	( $\mu$ , $r$ )
04/6/1-8/31	(3, 0.5)	05/6/1-8/31	(3, 1)	06/6/1-8/31	(3, 0.5)
04/9/1-11/30	(3, 1)	05/9/1-11/30	(3, 2)	06/9/1-11/30	(2, 0.5)
04/12/1-06/2/28	(3, 0.5)	05/12/1-06/2/28	(3, 4)	06/12/1-07/2/28	(3, 0.5)
05/3/1-5/31	(3, 2)	06/3/1-5/31	(3, 1)	07/3/1-5/31	(3, 1)
データ期間	( $\mu$ , $r$ )	データ期間	( $\mu$ , $r$ )	データ期間	( $\mu$ , $r$ )
07/6/1-8/31	(2, 0.5)	08/6/1-8/31	(3, 1)	09/6/1-8/31	null
07/9/1-11/30	(3, 1)	08/9/1-11/30	(2, 0.5)	09/9/1-11/30	null
07/12/1-08/2/28	null	08/12/1-09/2/28	(2, 0.5)	09/12/1-10/2/28	null
08/3/1-5/31	(1, 2)	09/3/1-5/31	(3, 4)	10/3/1-5/31	null
データ期間	( $\mu$ , $r$ )				
10/6/1-8/31	null				
10/9/1-11/30	(3, 1)				
10/12/1-11/2/28	(1, 4)				
11/3/1-5/31	(3, 4)				

パターンと範囲を広げて解析した。得られた延べ 40 の ( $\mu$ ,  $r$ ) の組み合わせの度数分布を図 6 に示す。興味深いのは、データが大きく 2 つのグループに分かれることである。すなわち  $\mu$ ,  $r$  とともに小さめのグループと、 $\mu$  が大きめで  $r$  が広い範囲で分布するグループである。

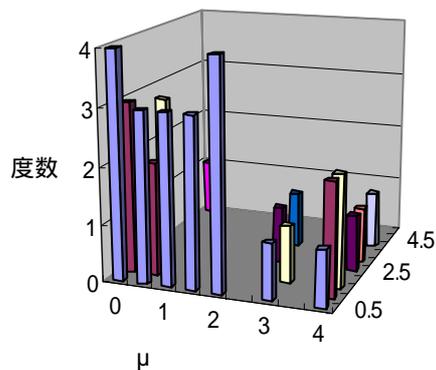


図 6 . 消費カロリーと体脂肪率の時系列データ間の相関係数が最大となる ( $\mu$ ,  $r$ ) セットの度数分布 (被験者 20 名、各 2 セット計 40 セット)

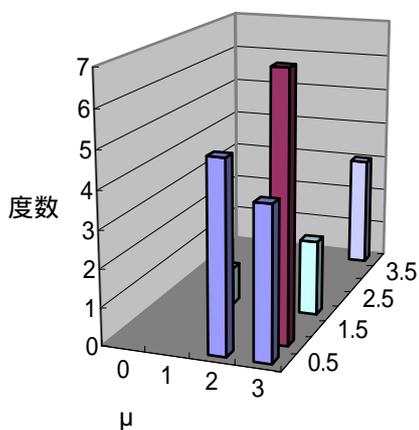


図 5 . 摂取カロリーと体脂肪率との時系列データ間の相関係数が最大となる ( $\mu$ ,  $r$ ) のセットの度数分布 (被験者 1 名、7 年間計 22 セット)

### 3.2. 複数被験者のデータ解析

20 人のボランティアの夏 (6/1~8/31) および秋 (9/1~11/30) の、体脂肪率と消費カロリーの時系列データにつき、3.1 と同様な解析を行った。ただし、個体差があることを想定し、 $\mu$  については、0~4 まで 0.5 刻みで、 $r$  については 0.5~4 まで 0.5 刻みとして、72

## 4. 結果の検討

運動や食事摂取等生活習慣データと体脂肪率等健康データの時系列相関解析において、相関係数の絶対値が最大となる重み付パラメータ ( $\mu$ ,  $r$ ) の組み合わせが、健康に関する個人の特質を表すという想定のもとに研究を行っている。今回、1 被験者の 7 年余りに亘る期間の時系列データをもとに、3 カ月区切り (季節毎) で ( $\mu$ ,  $r$ ) を求めたが、当初予想していた加齢や季節による変動は見られず、これらに関係なくばらついた値が得られた。消費カロリーおよび摂取カロリーと体脂肪率の相関に焦点を当てているが、体脂肪率の変動にはこれら以外の様々な生活習慣も関わっており、7 年余りの間には多くの因子がこれらの相関に影響を与えていると考えられる。したがって、得られた ( $\mu$ ,  $r$ ) の値がばらつくのは当然である。しかし、図 4 と図 5 から判るように、( $\mu$ ,  $r$ ) の特定の組み合わせの周辺で頻度が高くなる傾向がある。特に、体脂肪率と摂取カロリーの相関においては頻度が顕著に局在化している (図 5)。加齢や季節に関係がないということも勘案すると、これは個人の特質を表している可能性がある。すなわち、この被験者は体脂肪率と消費カロリーの関係においては (1~2, 0.5~1)、体脂肪率と摂取カロリーの関係においては (2~3, 0.5~1) が特徴という仮説に立つことになる。この被験者について

は、さらに今後傾向を追跡する予定である。

一方20名のボランティアの夏、秋半年間の体脂肪率と消費カロリーの関係を見ると、 $\mu = 0\sim 2$ で  $\mu$  が小さい ( $0.5\sim 1$ ) グループと  $\mu = 3\sim 4$  で  $\mu$  が広い範囲 ( $0.5\sim 4$ ) に分布するグループに分かれる。この場合も夏、秋の季節が影響を与えていることはなかったため、ここで立てられる仮説は、体脂肪率と消費エネルギーの関係において、比較的短期間のエネルギー消費があまり遅延なく ( $\mu = 0\sim 2$ ) 体脂肪率に影響を与える低  $\mu$  グループと、エネルギー消費が遅延をもって ( $\mu = 3\sim 4$ ) 体脂肪率に影響を与える高  $\mu$  グループに、クラスタリングされるということである。後者では、体脂肪率に影響を与えるエネルギー消費蓄積期間は広く分布する ( $\mu = 0.5\sim 4$ )。未だまだ被験者数が少ないので、これはあくまでも仮説であり、今後被験者数を増やすことによりさらに研究を進めていく。

## 5. まとめ

先の研究で、生活習慣データの時系列加算処理に  $\mu$  と  $\sigma$  の2つのパラメータを持つ正規分布関数型を活用した重み付けの概念を導入した。本報では、健康状態の変化と生活習慣データの時系列加算の間に最大の相関をもたらす重み付けのパラメータの組み合わせ ( $\mu, \sigma$ ) が個人の特質を反映しているという想定の下にデータ解析を行った。

その結果を以下にまとめる。

(1) 1被験者の7年余りに亘る体脂肪率、消費カロリーおよび摂取カロリーの時系列データを3カ月毎に区切って解析した結果、体脂肪率と消費カロリー、摂取カロリーそれぞれの関係において最大の相関係数を示す ( $\mu, \sigma$ ) の組み合わせが特定の組み合わせ範囲に偏っている。

(2) その偏りは、加齢や季節に関係がないことから、個人の特質を反映している可能性がある。

(3) 20被験者の半年間の体脂肪率と消費カロリーの時系列データを3カ月毎に区切って解析した結果、低  $\mu$  ( $\mu$  も小さい) グループと高  $\mu$  ( $\mu$  は広く分布) グループに分かれる傾向がみられる。

データ数からみてこれらの結果は未だ仮説に過ぎず、今後さらに研究を継続する。

## 謝辞

本研究は文部科学省科研費 (課題番号: 23500813) の助成を受けている。また、日本データベース学会と日立製作所による日立 HiRDB アカデミック制度の適用を受けている。

## 参考文献

- [1] H. Kumpusch, D. Hayn, K. Kreiner, M. Falgenhauer, J. Mor, and G. Schreier, "A Mobile Phone Based Telemonitoring Concept for the Simultaneous Acquisition of Biosignals and Physiological Parameters", Proc. 13<sup>th</sup> World Congress on Medical and Health Informatics (Medinfo2010), pp. 1344-1348, 2010.
- [2] E. C. Kyriacou, C. S. Pattichis, and M. S. Pattichis, "An Overview of Recent Health Care Support System for eEmergency and mHealth Applications", Proc. 31<sup>st</sup> Annual International Conference of the IEEE EMBS, pp.1246-1249, 2009.
- [3] H. Takeuchi, T. Hashiguchi, and T. Shintani, "Personal Dynamic Healthcare System Utilizing Mobile Phone and Web Technologies", Proc. 2<sup>nd</sup> Int'l Conf. Advances in Biomedical Signal and Information Processing, pp.304-307, 2004.
- [4] H. Takeuchi, Y. Mayuzumi, and N. Kodama, "Analysis of Time-Series Correlation Between Weighted Lifestyle Data and Health Data", Proc. 33<sup>rd</sup> Annual International Conference of the IEEE EMBS, pp.1511-1514, 2011.
- [5] 竹内裕之、児玉直樹, "生活習慣と健康状態に関する時系列データ解析手法の開発", DEWS 2008, E1-5, 2008.
- [6] H. Takeuchi, Y. Ikeda, and N. Kodama, "Time-Series Data Analyses for Healthcare-Data-Mining Based on a Personal Dynamic Healthcare System", Proc. 12<sup>th</sup> World Congress on Medical Informatics (Medinfo 2007), P309, 2007.