Disponible en ligne sur
ScienceDirect
www.sciencedirect.com
Elsevier Masson France
EM|consulte
www.em-consulte.com
IRBM
IRBM 35 (2014) 310–320

Healthcom 2013

Temporal progress model of metabolic syndrome for clinical decision support system

S. Jeong [a,*], C.-H. Youn [b], Y.-W. Kim [a], S.-O. Shim [c]

[a] *ETRI/PEC, 218 Gajeongno, Yusung-gu, Daejeon, 305-700, Republic of Korea*
[b] *KAIST/Dept. of Electrical Eng., 291 Daehak-ro, Yusung-gu, Daejeon, 305-701, Republic of Korea*
[c] *Department of Business and Accounting, Hanbat National Univ., Daejeon, Republic of Korea*

Received 13 January 2014; received in revised form 21 May 2014; accepted 26 August 2014

Available online 3 October 2014

Abstract

The development of an integrated and personalized healthcare system is becoming an important issue in the modern healthcare industry. One of main objectives of integrated healthcare system is to effectively manage patients having chronic diseases that require long term care and its temporal information plays an important role to manage the statuses of diseases. Thus, a patient having chronic disease needs to visit the hospital periodically, which generates large volume of medical examination data. Among the various chronic diseases, metabolic syndrome (MS) has become a popular chronic disease in many countries. There have been efforts to develop an MS risk quantification and prediction model and to integrate it into personalized healthcare system, so as to predict the risk of having MS in the future. However, the development of methods for temporal progress management of metabolic syndrome has not been widely investigated. This paper proposes a method for identifying the temporal progress of MS patients' status based on the chronological clustering methodology. To investigate the temporal changes of disease status, we develop a chronological distance variance model that quantifies the difference of areal similarity degree (ASD) values between estimated and examined results of MS risk factors. We evaluate the clinical effectiveness of the temporal progress model by using sample subjects' examination results that have been measured for 10 years. We further elaborate the accuracy of the proposed temporal progress estimation method by using multiple linear regression method. Then, we develop a tier-based patients' MS status classification based on the chronological distance variance. The tier classification is based on the sensitivity for temporal change of MS status according to different values of control range of chronological distance variance. Our proposed temporal change identification method and patients' tier classification are expected to be incorporated with the integrated healthcare systems to help physicians with identifying the temporal progress of MS patients' health status and MS patients with self-management at home environments.

1. Introduction

The development of an integrated and personalized healthcare system is becoming an important issue in the modern healthcare industry. One of main objectives of integrated healthcare system is to effectively manage patients having chronic disease. Among the various chronic diseases, metabolic syndrome (MS) has become a major public healthcare issue in many countries. The importance of MS is that it is related to the occurrence of major chronic diseases such as diabetes and cardiovascular disease. Thus, early detection of the risk for having MS can help to proactively manage the major chronic diseases. There have been efforts to develop MS risk quantification and prediction model and to integrate them into personalized healthcare system, so as to predict the risk of having metabolic syndrome in the future [1]. However, the development of methods for temporal progress management of MS has not been widely investigated. Different from acute disease, chronic disease requires long term care and its temporal information plays

* Corresponding author.
E-mail addresses: sjjeong@etri.re.kr (S. Jeong), chyoun@kaist.ac.kr (C.-H. Youn), qkim@etri.re.kr (Y.-W. Kim), soshim@hanbat.ac.kr (S.-O. Shim).

http://dx.doi.org/10.1016/j.irbm.2014.08.003
1959-0318/

an important role to manage the status of disease. Since patients having chronic diseases such as MS typically spend most of time out of hospital environment between regular health examinations, providing the patients' health status over time will be helpful for patients to perform self-management at home environment.

Furthermore, when changes in patients' disease status occur, it is important to decide whether the changes are within control ranges of patients' status. Especially, patients having chronic disease routinely visit hospital and spend most of time in out-of-hospital environments. Thus, it is helpful for a patient to perform self-management by providing a method for identifying the patient's disease changes are within a designated control range. Also, during at home disease care, patients may generate large volume of medical data and physicians need to keep tracking the patients' status over time including routine medical examinations. However, it is difficult for physicians to review the huge number of medical information during the patient's hospital visit due to limited time for consulting. Thus, the results from these examinations and patients need to be classified according to patients' disease status to facilitate for physicians to effectively care chronic disease patients during their hospital visits.

To resolve these issues in this paper, we firstly investigate the problem spaces regarding temporal management and patient disease status. Then, we propose a method for identifying a temporal progress and patient's status of MS. Further, the effectiveness of the proposed method is evaluated using a sample patient data while emphasizing the capability to identify chronological changes of MS status. The proposed method supports temporal progress care of MS status using a chronological distance analysis on the areal similarity degree (ASD) [1,2]. The proposed method can be used to advance MS care services in which it combines the MS diagnosis services and temporal progress care services. Among chronic diseases, this paper targets MS because the criteria for diagnosing MS consist of thresholds based on scalar values and the quantification model of the risk and medical report are investigated in several studies, and the development of MS can impact on the development of more serious chronic diseases [1–3,12]. Therefore, we develop a temporal progress model of MS. After that, to efficiently manage patients with MS, we propose a patient status classification scheme for MS care based on the quantified changes of disease status. As discussed in [3], an individual having a chronic disease such as MS typically spends most of time out-of-hospital environments and the self-management period plays an important role for caring the chronic diseases. Thus, it is necessary to develop a method for delivering self-management results to the hospital healthcare system and for helping the physician investigate the temporal progress during home caring. We expect the proposed temporal progress method and patient tier classification scheme to be utilized for investigating the at-home management of MS. Finally, we summarize the contributions of this paper and list the limitations of proposed method and further study issues.

2. Problem description

Different from acute disease, chronic disease requires long term care and its temporal information plays an important role to manage the status of disease. Several issues and problems of conventional healthcare system have been reported in [3]. Among them, difficulty in identifying important information from large datasets containing diverse types of information in a laboratory or clinical situation becomes a problem among physicians and clinicians. It is also important to find out whether a patient's disease status is under control and within the normal range. To resolve the limitation, various types of health information systems have been developed to support the physicians for caring patients in terms of temporal disease progress management. However, most of the health information systems are focused on visualizing temporal data in a timeline, while displaying specific events from the patient data. The overall goal of these health information systems is to present physicians temporal information contained in a record, improving their ability to recognize patterns for knowledge discovery and following treatment. They introduce simple visualization tools, but some include automated computational enhancements supporting it. However, due to the complexity of the underlying data, a lot of further work is required to address these intricacies rather than using simplistic approaches [4–6].

Advances in patient caring and monitoring technologies have allowed physicians to track a patient's physiological state more closely and more accurately. These technologies enable out-of-hospital health monitoring. With the increasing amount of electronic medical data, system assisted medical decision should be adopted to effectively provide health care services. There are several types of systems for providing medical services and healthcare services such as health information management system (HIMS), clinical decision support system (CDSS), hospital information system (HIS), and so on. HIMS is being used to support key medical care procedures, to make medical decision and prescription, to manage patient's health conditions, or even for hospital administration, respectively. Health information management is the practice of acquiring, analyzing and protecting digital and traditional medical information vital to providing quality patient care. CDSS is interactive expert system, which is designed to assist physicians and other health professionals with decision making tasks, such as determining diagnosis of patient data. HIS is a comprehensive, integrated information system designed to manage all the aspects of a hospital operation, such as medical, administrative, financial, legal, and the corresponding service processing [13]. The traditional CDSS has been developed to investigate patient examination results and to provide expert advice for managing a patient's specific condition. Furthermore, patients having chronic disease need to visit the hospital periodically, which generates large volume of medical data. Physicians need to keep tracking the patients' status over time including routine medical examinations. It is inadequate for physicians to review the huge number of medical information for chronic patients. Thus, the results from these examinations and patients need to be classified according to patients' disease status and the changes of

disease status need to be determined whether the disease status changes are under control [7].

3. Proposed methodology for identifying temporal progress of MS

3.1. Chronological clustering method for identifying temporal variation of MS

Chronological clustering is a punctuation equilibrium model that is learning the sequence within biological communities. It is used to investigate the temporal variation of observations or data by analyzing the similarity between them. There are two required elements to perform chronological clustering method, namely connectedness and the sensitivity level. Connectedness means that the data or observations of data are related to each other as well as they are distributed over time domain. Sensitivity level denoted by α, is the clustering tuning indicator. This indicator is the primary element in defining the similarity between adjacent observations. The indicator α is the tolerance ratio of similarity or correlation between observations [8,9]. The chronological clustering method has been applied to investigate the temporal progress of chronic diseases [10]. A chronic disease patient requires regular medical examination even though no noticeable symptom is observed. Thus, if two consecutive examination results show a slight variation between them, a physician may ignore the second examination results as it is a repetition of the previous one. Also, in that case, the patient may skip regular visits to a hospital to see a physician, which can save time and reduce the cost to the patient. To support this scenario, Hazemi et al. [10] developed a distance function and a sensitivity of change function as follows:

Definition 1. A medical report is a set of medical examination results containing one or more examination criteria represented as scalar values. A normalized report value indicates that a single integrated number of one or more examination results consisting the medical report, and the range of the single integrated number is from 0 to 1. Previous report and current report are chronologically ordered medical reports examined at t_1 and t_2, where $t_1 < t_2$ holds.

Definition 2. Given two consecutive normalized medical report values containing N examination criteria, the distance function $d_N(2)$ is defined as:

$$d_N(2) = \frac{\sum_{i=1}^{N} |\text{Previous normalized medical report value} - \text{Current normalized medical report value}|}{\text{Number of examination criteria}} \quad (1)$$

Let the previous normalized report value and the current normalized report value be VP_i and VC_i, respectively and N denote the number of examination criteria contained in a medical report. The ranges of VP_i and VC_i are from 0 to 1. For example, if a medical report contains two examination criteria such as blood pressure and weight, N becomes 2. Then $d_N(2)$ can be expressed as

$$d_N(2) = \frac{\sum_{i=1}^{2} |VC_i - VP_i|}{2} \quad (2)$$

Definition 3. Given the sensitivity level denoted by α and distance function d_N, the importance of information (IoF) for two normalized medical reports containing N examination criteria each is defined as:

$$IoF = \begin{cases} 0, & \text{if } d_N(2) \le 1-\alpha. \text{ NOT noticeable change} \\ & \xrightarrow{\text{yields}} \text{Discard it} \\ 1, & \text{if } d_N(2) > 1-\alpha. \text{ Noticeable change} \\ & \xrightarrow{\text{yields}} \text{Visualize it} \end{cases} \quad (3)$$

The value of $d_N(2)$ in Eq. (2) is related to the indicator α that is the sensitivity level of the change function. When IoF is zero, which means that $d_N(2)$ is less than or equal to $(1-\alpha)$, the changes in VC_i in comparison with VP_i are small and the patient's status has not noticeably changed. Thus, the current report may not need to be checked by the physician. If $d_N(2)$ is larger than $(1-\alpha)$, it means that the VC_i has changed noticeably and the current report must be investigated by the physician. Since IoF relies on the value of α, it is important to decide the value of α that is clinically meaningful. The range of α is from 0 to 1, where the correlation between the current and previous report becomes weaker as α becomes closer to 0. The correlation becomes stronger as α becomes closer to 1. A weak correlation, that is, when α is close to 0, causes the threshold for noticeable change to be high. Thus, the current report is discarded unless it shows significant difference from the previous one. For example, if α is 0.1, then the reports are considered similar whenever there is more than 10% similarity between the two reports. Also, if α is 0.8, then the reports considered to be similar whenever there is more than 80% similarity between them [10]. Therefore, the effectiveness of the method is largely dependent on the determination of α.

However, the authors of [10] did not provide any further information on determining the value of α or quantitatively integrating the N examination results contained in a medical report into single integrated quantitative value, which are the critical factors to utilize the chronological clustering method. Since the statuses of diseases are generally described by multiple examination criteria, it is essential to map the examination results of criteria in a medical report on scalar scale between 0 and 1.

To resolve the limitations of the above approach, we propose an ASD-based chronological clustering method to identify the temporal progress of chronic disease, especially MS. ASD is a similarity analysis model between two weighted radar charts comprising MS diagnostic criteria and examination results of risk factors, which was proposed by Jeong et al. [1,11]. The authors in [1,11] utilized ASD to establish risk quantification model for MS, which quantifies the disease status as a number between 0 and 1, and they determined that the risk model could effectively represent the disease status. The model resolves the well-known issues of MS diagnostic criteria as follows [12]:

- Equally weighted MS diagnostic criteria
- No explanation for disease status and changes of status

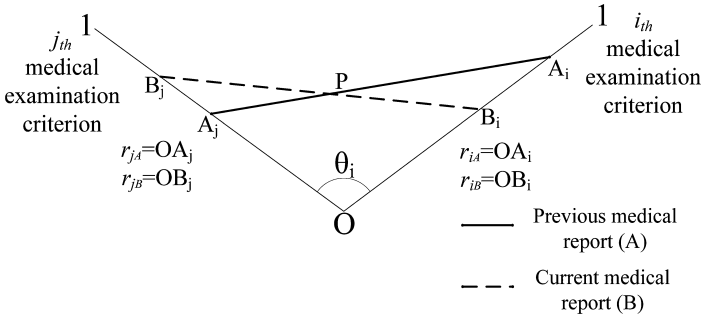

Fig. 1. A partial radar chart consisting of two polygons. The figure has been adopted from [1] and modified.

– No information about quantified risk by comparing examination results with diagnostic criteria

Hence, the ASD-based risk model can be applied to identify the temporal change of chronic disease, particularly MS. Let us consider the partial weighted radar chart shown in Fig. 1.

Definition 4. According to [1], the ASD of a partial weighted radar chart comprising two polygons A_{ij}, i.e., $\Delta O A_i A_j$, and B_{ij}, i.e., $\Delta O B_i B_j$, is defined as

$$ASD = S(A_{ij}|B_{ij}) = \frac{\text{Area of intersection of polygons } A_{ij} \text{ and } B_{ij}}{\text{Area of polygon } B_{ij}} \quad (4)$$

Theorem 1. *The distance between two chronologically ordered medical reports expressed by a partial weighted radar chart containing two examination criteria in each report can be calculated as:*

$$d_2(2) = 2 - ASD(A_{ij}|B_{ij}) - ASD(B_{ij}|A_{ij}), \quad (5)$$

where A_{ij} and B_{ij} are partial weighted radar charts constructed using the normalized values of two medical examination criteria, namely i and j in each medical report and $j = (i+1) \bmod 2$.

Proof. Let us assume that two medical reports A_{ij} and B_{ij} are results of the examination criteria of the current and previous medical reports expressed by a partial weighted radar chart. Then, according to Eq. (2), $d_2(2)$ becomes

$$d_2(2) = \frac{\sum_{i=1}^{2} |A_{ij} - B_{ij}|}{2} = \frac{|A_{1,2} - B_{1,2}| + |A_{2,1} - B_{2,1}|}{2} = \frac{|A_{ij} - B_{ij}| + |A_{ji} - B_{ji}|}{2} \quad (6)$$

According to Fig. 1, $|A_{ij} - B_{ij}|$ equals to $|A_{ji} - B_{ji}|$. Thus, Eq. (6) becomes $|A_{ij} - B_{ij}|$. Since $|A_{ij} - B_{ij}|$ indicates the normalized difference between A_{ij} and B_{ij}, and A_{ij} and B_{ij} can be mapped into a scalar value by means of their areas. Therefore, it becomes

$$|A_{ij} - B_{ij}| = \frac{\text{Area of } A_{ij} - \text{Intersection of } A_{ij} \cap B_{ij}}{\text{Area of } A_{ij}} + \frac{\text{Area of } B_{ij} - \text{Intersection of } A_{ij} \cap B_{ij}}{\text{Area of } B_{ij}} \quad (7)$$

According to Definition 4, since the ASD of a partial radar chart is defined as the ratio of the intersection of A_{ij} and B_{ij} over B_{ij}, Eq. (7) becomes

$$|A_{ij} - B_{ij}| = \{1 - S(A_{ij}|B_{ij})\} + \{1 - S(B_{ij}|A_{ij})\} = 2 - S(A_{ij}|B_{ij}) - S(B_{ij}|A_{ij})$$

Therefore, we can obtain

$$d_2(2) = 2 - ASD(A_{ij}|B_{ij}) - ASD(B_{ij}|A_{ij}). \quad \square \quad (8)$$

The conventional MS diagnosis criteria only determine whether a subject is diagnosed as having MS. However, chronic diseases such as MS that are characterized by temporal progress and periodic examinations need a method to easily determine whether the subject's disease status has changed since the previous examinations. The distance in Theorem 1 makes it easy to test whether the examined subject's disease status has changed over the patient specific sensitivity level described in Definitions 2 and 3. Further, the distance function can be calculated using the ASD of the subject's examination results and the thresholds of MS diagnosis criteria; thus, the potential risk of having MS can be easily quantified [1]. Therefore, Theorem 1 can be utilized as a preliminary diagnosis method in clinical decision support system for chronic disease.

Corollary 1. *Let $d_N(2)$ be the distance of two chronologically ordered medical reports containing N examination criteria in each. According to Theorem 1, $d_N(2)$ can be calculated as:*

$$d_N(2) = 2 - ASD(R_A|R_B) - ASD(R_B|R_A), \quad (9)$$

where R_A and R_B are weighted radar charts constructed using N disjoint partial weighted radar charts, $R_A = \{A_{12}, A_{23}, \ldots, A_{n1}\}$ and $R_B = \{B_{12}, B_{23}, \ldots, B_{n1}\}$. Each weighted radar chart represents a medical report with N medical examination criteria.

Since a weighted radar chart is a sum of disjoint partial weighted radar charts and Theorem 1 provides the calculation of distance function for two partial weighted radar charts, Corollary 1 is obvious.

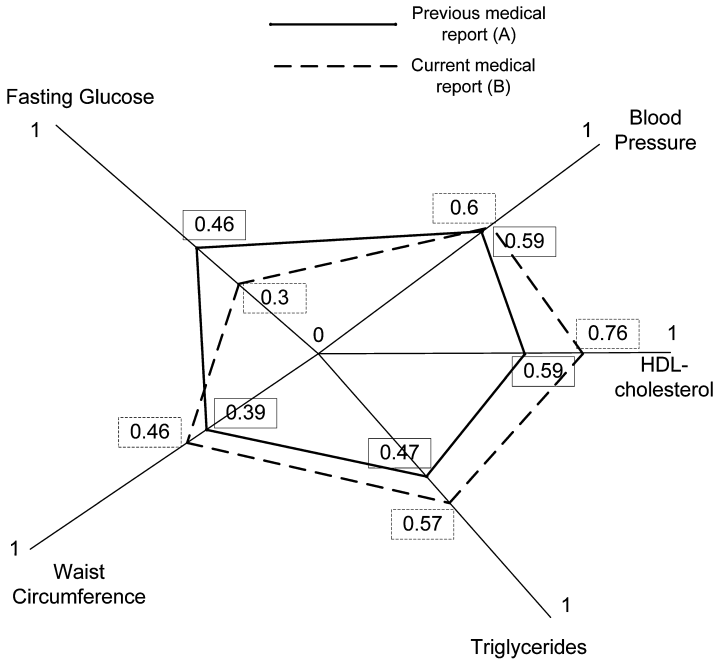

Fig. 2. Two weighted radar charts constructed from two data sets A and B. The figure has been adopted from [1] and modified.

According to Eq. (9), we can calculate the distance function $d_5(2)$ for two weighted radar charts containing 5 examination criteria shown in Fig. 2 as 0.31.

4. Evaluation and discussion

4.1. Evaluation and discussion of MS temporal progress identification method

This section presents a method to determine a patient's temporal disease progress based on chronological distance described in Corollary 1. Since it had been investigated that ASD can be used to quantifying and identifying risk of MS [1], we can assume that the variation of ASD over time, which is defined as a chronological distance in this section, indicates the change of the risk status of MS. Therefore, we propose a method for identifying the temporal disease status change by analyzing the chronological distances.

Hypothesis 1. If ASD quantifies the risk of having a chronic disease, particularly metabolic syndrome, we can claim that the chronological distance, i.e., $d_N(2)$, describing the difference between ASD value and linear regression value of ASD values over time can be used to identify the disease status change of a patient.

To verify the hypothesis, we perform linear regression analysis of ASD values over time. Let $y(T)$ be a linear regression of ASD values from t_0 to t_i. Then, $y(T)$ can be expressed as follows:

$$y(T) = aT + b \tag{10}$$

Let $y(t_i)$ and $y(t_{i+1})$ be regression result at time t_i and t_{i+1} respectively. Also, let $y'(t_{i+1})$ be examined ASD value at time t_{i+1}. Then, we can define the chronological distance variance as follows:

Definition 5. Chronological distance variance over t_i and t_{i+1} is defined as

$$\left|y(t_i) - y'(t_{i+1})\right| \tag{11}$$

Let δ be the acceptable range of variation, then we claim that the chronological distance variance is within a control range if

$$\left|y(t_i) - y'(t_{i+1})\right| < y(t_{i+1}) * \delta \tag{12}$$

holds.

In other words, if the examined chronological distance at time t_{i+1} is within $\delta\%$ range from the regression value at time t_{i+1}, the disease status change at from t_i to t_{i+1} is within the control range. If the status change exceeds the control range, the patient will be instructed to be diagnosed in detail. By considering Eq. (10) and Eq. (11), we can obtain the following observations. In Eq. (10), the slope a implies the temporal trend of subject's disease status. The values of a is defined as follows:

$a \approx 0$ implies that the subject's disease status does not changed since last medical examination.

$a > 0$ implies that the subject's examination results for criteria is in increasing trend.

$a < 0$ implies that the subject's examination results for criteria is in decreasing trend.

In Eq. (11), the chronological distance variance quantifies the integrated change of examination results for disease criteria.

$|y(t_i) - y'(t_{i+1})| < y(t_{i+1}) * \delta$ implies the subject's disease change is within a control range.

$|y(t_i) - y'(t_{i+1})| \geq y(t_{i+1}) * \delta$ implies the subject's disease change is outside a control range.

Table 1
A sample patient's medical examination results for metabolic syndrome risk factors.

Examination time	T0	T1	T2	T3	T4	T5	T6	T7	T8
BMI (kg/m^2)	23.6	23.5	22.4	22.6	21.7	22.6	21.7	21.4	21.9
Fasting glucose (mg/dl)	92	91	96	89	89	87	94	85	91
Waist circumference (cm)	80	77	82	83	78	77	79	82	77
HDL-cholesterol (mg/dl)	58	49	61	47	58	42	61	63	57
Triglycerides(mg/dl)	120	97	84	91	70	87	83	78	94
Systolic BP (mm Hg)	116	120	110	110	110	100	120	110	114
Diastolic BP (mm Hg)	64	80	70	70	70	60	60	70	75

BMI, body mass index; BP, blood pressure.

Table 2
Calculation of chronological distance of a sample patient.

Time	$t_0 - t_1$	$t_1 - t_2$	$t_2 - t_3$	$t_3 - t_4$	$t_4 - t_5$	$t_5 - t_6$	$t_6 - t_7$	$t_7 - t_8$
Chronological distance	0.042	0.064	0.021	0.110	0.034	0.001	0.062	0.063

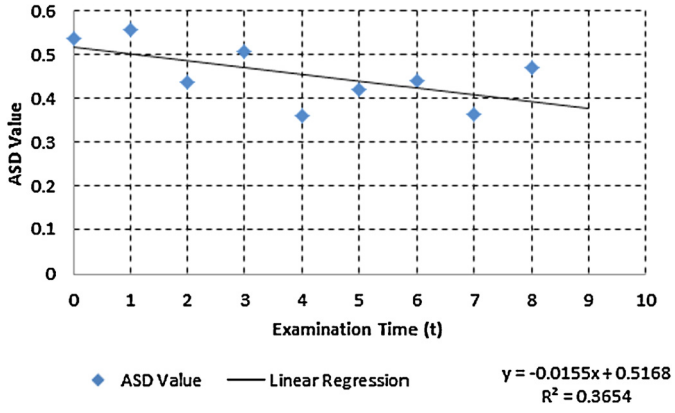

Fig. 3. Linear regression of ASD values over time.

To verify the proposed Hypothesis 1, we performed an evaluation with a sample patient's medical examination results. Table 1 shows the characteristics of the study subject over time.

Then, we obtained the linear regression result as shown in Fig. 3. The regression analysis resulted in the following statistics. The *R-squared* value is 0.3654 and the analysis showed 95% confidence level with 0.0598 error. Also, the top 5% and the bottom 5% percentile showed 0.6350 and 0.4296, respectively.

The calculated chronological distances using Table 1 and Eq. (10) are shown in Table 2.

Let the control range at time t_{i+1} be $A(t_{i+1})$. Also, let us assume that the control range parameter δ be 0.25 (or 25%), then we can get the control range of chronological distance at time t_{i+1}, as follows:

$$0.2946 < A(t_{i+1}) < 0.4910 \tag{13}$$

Since ASD value at time t_{i+1} is 0.4708 and within the control range $A(t_{i+1})$, we can decide that the sample patient's disease status is within the control range. Therefore, when the examined chronological distance at time t_{i+1} is outside the range shown in Eq. (13), the subject is likely to have major change at disease status and needs to consult physician. The analysis in this sub-section has been performed with very limited sample examination results. Thus, more extended clinical evaluation with a large number of patient examination data is necessary in order to establish clinically meaningful value of δ. Also, longitudinal study using cohort is need to extensively evaluate the clinical effectiveness of the proposed temporal progress identification method.

4.2. ASD value estimation based on multiple regressions

To effectively manage temporal progress of MS, it is important to identify the temporal trend of a subject's health status comprising MS risk factors. In the previous section, we developed a chronological clustering model-based chronological distance variance method. The health status of the subject can be controlled and monitored by using the chronological distance variance and control range presented in Eq. (12). Since ASD quantifies the risk of MS for a subject, it is possible to observe the health status changes over time by estimating the future ASD value of the subject. Since our proposed temporal progress model utilizes an estimated ASD value for observe the temporal change, the effectiveness of the proposed temporal model relies on the accuracy of the estimated future ASD values. Since the future ASD value estimation in Eq. (12) uses a linear regression method, we further improve the accuracy of estimating future ASD values utilizing multiple linear regressions in this section. To analyze the effectiveness of linear regression of ASD values in estimating future ASD values, we performed longitudinal analysis using two sample subjects. Tables 3 and 4 show the medical examination results of the two subjects.

We performed linear regression as ASD values using examination results from T1 to 2011 (T8 for Subject 1 and T9 for Subject 2). Then, we compared the actual ASD examination results in 2012 with estimated ASD results using regression function. Fig. 4 shows ASD estimation result of Subject 1. The linear regression of ASD values for Subject 1 shows R value as 0.7812, which indicates linear relationship between time and ASD values. The linear regression function of Subject 1 is as

$$y(t) = -0.0238x(t) + 0.5602, \qquad R = 0.7812. \tag{14}$$

Table 3
Medical examination results of Subject 1 for retrospective clinical analysis (Young male; ages <40 years).

Exam. time	T1 (2004)	T2 (2005)	T3 (2006)	T4 (2007)	T5 (2008)	T6 (2009)	T7 (2010)	T8 (2011)	T9 (2012)
Age (years)	27	28	29	30	31	32	33	34	35
FG (mg/dl)	92	91	96	89	89	87	94	85	91
WC (cm)	80	77	82	83	78	77	79	82	77
HDL-C (mg/dl)	58	49	61	47	58	42	61	63	57
TG (mg/dl)	120	97	84	91	70	87	83	78	94
SBP (mm Hg)	116	120	110	110	110	100	120	110	114
DBP (mm Hg)	64	80	70	70	70	60	60	70	75

Table 4
Medical examination results of Subject 2 for retrospective clinical analysis (Middle-aged male; 40≤ ages <65 years).

Exam. time	T1 (2003)	T2 (2004)	T3 (2005)	T4 (2006)	T5 (2007)	T6 (2008)	T7 (2009)	T8 (2010)	T9 (2011)	T10 (2012)
Age (years)	46	47	48	49	50	51	52	53	54	55
FG (mg/dl)	99	108	105	130	116	147	138	148	129	173
WC (cm)	95	95	95	95	94	96	98	94	97	88
HDL-C (mg/dl)	50	61	42.4	53.4	35.5	50.4	46.5	51.7	53.6	54.4
TG (mg/dl)	190	137	216	153	250	133	206	125	101	172
SBP (mm Hg)	143	140	110	140	130	124	130	125	120	122
DBP (mm Hg)	94	100	90	90	100	82	90	90	85	76

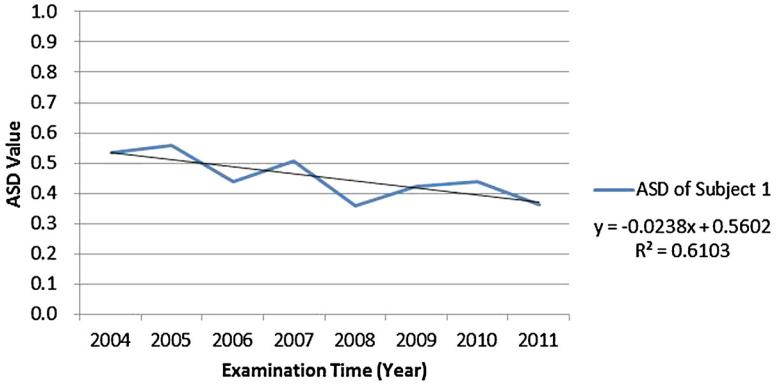

Fig. 4. Linear regression of ASD values of Subject 1.

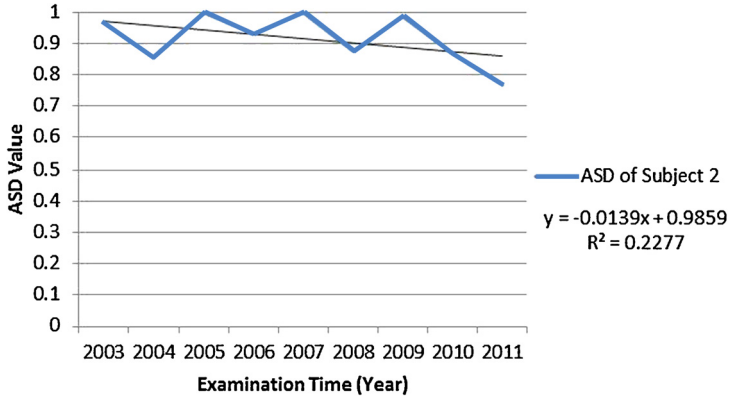

Fig. 5. Linear regression of ASD values of Subject 2.

Using Eq. (14), we can estimate ASD value at time 2012 as 0.322. The examined ASD value at time 2012 by using Table 3 can be calculated as 0.471. Thus,the error of estimation is calculated as 31.57%.

Similarly, we performed linear regression for Subject 2's ASD values over time and Fig. 5 shows the regression results. The linear function of Subject 2 is as

$$y(t) = -0.0139x(t) + 0.9859, \qquad R = 0.4771. \tag{15}$$

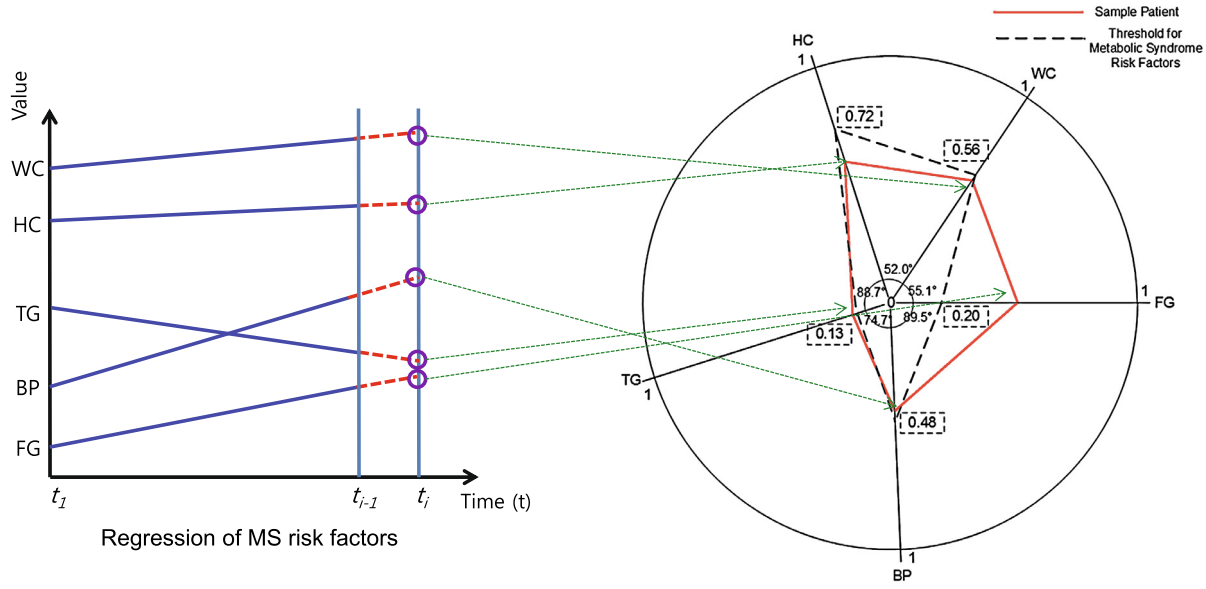

Fig. 6. Application of linear regressions of MS risk factors to future ASD estimation.

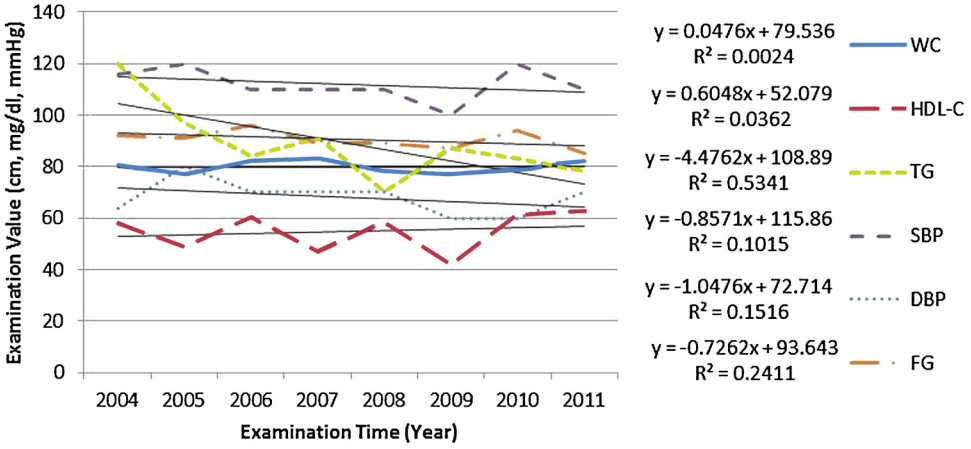

Fig. 7. Linear regression results for MS risk factors of Subject 1.

Using Eq. (15), we can estimate ASD value at time 2012 as 0.847. The examined ASD value at time 2012 by using Table 4 can be calculated as 0.902. Thus, the error of estimation is calculated as 6.10%.

To increase the accuracy of ASD value estimation, we performed multiple linear regressions for each MS risk factor's values from t_1 to t_{i-1}. Then, the regression functions for each risk factor are determined as follows:

$$y_{WC}(t_i) = a_{WC} \times x(t_i) + b_{WC}$$
$$y_{HC}(t_i) = a_{HC} \times x(t_i) + b_{HC}$$
$$y_{TG}(t_i) = a_{TG} \times x(t_i) + b_{TG}$$
$$y_{BP}(t_i) = a_{BP} \times x(t_i) + b_{BP}$$
$$y_{FG}(t_i) = a_{FG} \times x(t_i) + b_{FG} \quad (16)$$

where *WC*, Waist Circumference; *HC*, HDL-cholesterol; *TG*, Triglycerides; *BP*, Blood Pressure; *FG*, Fasting Glucose.

Since there exist five MS risk factors, we separately performed linear regression for each risk factor and the values of risk factors at time t_i are estimated using the regression function in Eq. (16). Fig. 6 shows the application of linear regressions of risk factors to future ASD estimation. Regression function of each risk factor is determined by using values from t_1 to t_{i-1}. Then, the estimation value at time t_i is calculated using the regression function. The estimated values of MS risk factors are utilized for calculating the ASD value at time t_i.

Fig. 7 and Fig. 8 show the linear regression results for Subject 1 and Subject 2's MS risk factors' examination values. From the analysis results, we can observe that some risk factors' temporal values are well fitted into linear function, but others do not show strong relationship between examination values and time. To compare the accuracy for estimating the future ASD value between ASD-based regression and MS risk factors-based regression, we calculated future ASD value using risk factors' estimated values.

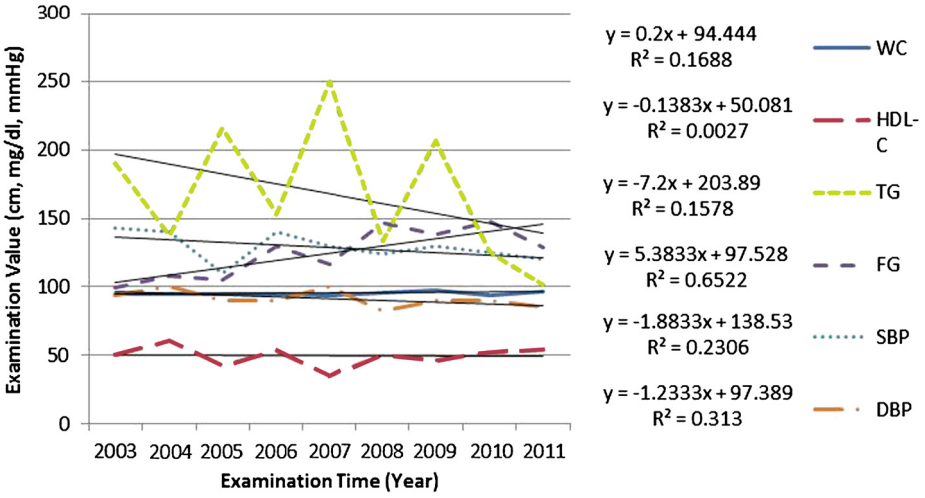

Fig. 8. Liner regression results for MS risk factors of Subject 2.

Table 5
Summary of ASD estimation accuracy comparison.

	Subject	ASD-based regression	Risk factors-based regression
Estimated ASD value at time t_i (2012)	Subject 1	0.322	0.350
	Subject 2	0.847	0.896
Examined ASD value at time t_i (2012)	Subject 1		0.471
	Subject 2		0.902
Estimation error (%)	Subject 1	31.57	25.72
	Subject 2	6.10	0.69

Table 5 shows the comparison of ASD value estimation accuracy between ASD-based regression and risk factors-based regression for the sample subjects. According to the table, we can observe that the MS risk factors-based regression method indicates better accuracy for estimating future ASD value than ASD-based regression method.

4.3. Determination of control range thresholds for patient tier classification based on chronological clustering

This section describes the determination of patient tier classification thresholds based the control range of chronological distance variance described in Eq. (12). In Eq. (12), it is determined that whether the MS status changes are within the control ranges of the patient by comparing the chronological distance variance δ with the difference between an estimated ASD value through linear regression and examined ASD value at a given time. If the difference between the two ASD values is less than δ, it is determined that the patient's disease status is under control. Therefore, to apply the Hypothesis 1 to identifying the temporal changes of a patient's MS status, it is necessary to develop reference values for δ. In this section, we determine the quantitative values for the references of δ.

Fig. 9 shows the classification criteria for representing the patients' MS status changes. In the figure, the green graph indicates that the current medical report has not been changed since the previous report. The blue, purple, and red graphs indicate that there are minor, major, and significant change between the estimated and examined ASD values, respectively. Thus, a physician can easily recognize the temporal change on patient's disease status with visualized interface. In the significant change status, the sensitivity level α is determined as TH_3 to indicate significant change of patient disease status. We chose $0.996\,(3\sigma)$ for the value of TH_3, so when the distance value $d_N(2)$ of the estimated and examined ASD values is greater than or equal to 0.996, it is immediately notified to the physician because the patient disease status was significantly changed since the previous examination. For the major change status, sensitivity level α is determined as TH_2 to indicate moderate and major change of patient disease status. We chose $0.954\,(2\sigma)$ for the value of TH_2. So, when $d_N(2)$ is between 0.954 and 0.996, the patient's status change is classified as major change. For the minor change status, the sensitivity level α is determined as TH_1 to indicate changes including minor change of patient disease status. We chose $0.682\,(1\sigma)$ for the value of TH_1. So, when $d_N(2)$ is between 0.682 and 0.954, the patient's disease status change is classified as minor change. When $d_N(2)$ is less than 0.682, the patient's MS status is considered to be within the control range.

According to the disease care service levels of MS shown in Fig. 9, we classify MS patients care services into four levels according to the patients' disease change status, the significant status for *TIER*(1), the major status for *TIER*(2), the minor status for *TIER*(3), and the controlled status for *TIER*(4), respectively. Table 6 lists the criteria for patients tier classification. We consider MS only in this section, so Table 6 describes criteria for MS patients. However, the criteria may be applicable to other chronic diseases, if risk quantification models for other chronic diseases are established. Each patient's tier is determined based on the proposed tier classification criteria. The very low risk patient is classified into *TIER*(1) group. A patient with low risk is classified into *TIER*(2) group, while medium or high risk patient is classified into *TIER*(3) and *TIER*(4), respectively.

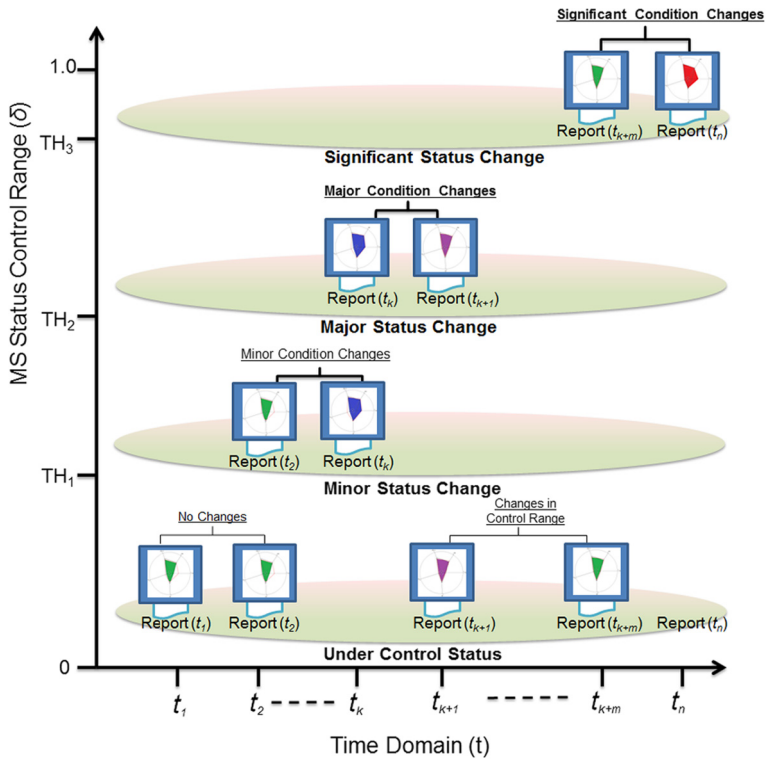

Fig. 9. Sensitivity for temporal change of MS status according to different values of δ. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Table 6
Criteria for determining thresholds of patients MS status changes.

Patient TIER	*TIER(1)* ($d_N(2) < TH_1$)	*TIER(2)* ($TH_1 \le d_N(2) < TH_2$)	*TIER(3)* ($TH_2 \le d_N(2) < TH_3$)	*TIER(4)* ($TH_3 \le d_N(2)$)
Criteria for TIER thresholds	Control range: $y(t_{i+1}) * \delta$			
	Less than 34.1% (1σ)	Less than 47.7% (2σ)	Less than 49.8% (3σ)	Greater than 49.8%

5. Conclusion

Because of its usefulness to identify individuals at high risk for CVD and T2DM, MS has become a major concern of many countries. There has been much effort to establish diagnostic criteria for MS, but it is known that current diagnostic criteria of MS have following weaknesses [1]:

- No consideration for different importance among risk factors
- Thresholds-based binary style diagnosis
- Difficulty in estimating the risk of MS for non-MS subjects
- Difficulty in managing the temporal change of the status of MS risk factors

Among the known weaknesses, to resolve the difficulty in managing the temporal changes of MS status, we proposed a novel method for identifying temporal changes of MS patients' status based on the chronological clustering method. To identify the temporal changes, we developed a chronological distance variance model that calculates the difference between estimated and examined ASD values. The clinical effectiveness of the proposed model was evaluated using sample patients data over 9 years and the evaluation results showed that the proposed model can be used to analyze the temporal change of MS status. Using the proposed model, we can effectively manage the patients having metabolic syndrome using the control range of patient status. Then, we developed the tier-based patients' MS status classification based on the chronological distance variance. The tier classification is based on the sensitivity for temporal change of MS status according to different values of control range of chronological distance variance. The evaluation results showed that the proposed patients tier classification could help physicians with identifying the major status changes of MS patients' risk factors. Our proposed temporal change identification method and patients tier classification are expected to be incorporated with the integrated healthcare systems to help physicians with identifying the temporal progress of MS patients' health status and MS patients with self-management at home environments. Also, by using the proposed method, we can design new application services, such as analyzing long-term trends of patient's disease status, a knowledge-based decision support tool for MS.

The limitations and recommendations for further study issues of this paper can be summarized as follows:

– The proposed method identifies the temporal progress of MS based on the chronological clustering method of ASD values. The clinical effectiveness of the proposed method is examined by using two subjects' 10-years longitudinal medical examination results. Our evaluation results have shown that the proposed method could estimate the temporal changes of MS status. However, to further test the validity of the temporal progress method, a longitudinal study should be performed to determine the clinically meaningful values of sensitivity level α and control range δ in a cohort of subjects.
– The two sample examination data consist of young male and middle-aged male subject, respectively. Further the young male subject was not diagnosed as MS, whereas the middle-aged male was diagnosed as MS. Our analysis results have shown the proposed temporal method is effective for both non-MS and MS subjects. However, the proposed method has not been evaluated for female subjects. Therefore, it is further required to evaluate the proposed method by using the longitudinal examination data of female subjects.

Acknowledgements

This research was supported by the ICT Standardization Program of MISP (The Ministry of Science, ICT & Future Planning). This research was supported by the Ministry of Science, ICT&Future Planning, Korea, under the ITRC support program (NIPA-2014(H0301-14-1020)) supervised by the National IT Industry Promotion Agency.

References

[1] Jeong S, Jo YM, Shim S-O, Choi Y-J, Youn C-H. A novel model for metabolic syndrome risk quantification based on areal similarity degree. IEEE Trans Biomed Eng 2014;61(3):665–79.
[2] Jeong S, Youn C-H, Kim Y-W. A method for identifying temporal progress of chronic disease using chronological clustering. In: IEEE int. conf. e-health networking, application & services. 2013. p. 314–8.
[3] Jeong S, Youn C-H, Shim EB, Kim M, Cho YM, Peng L. An integrated healthcare system for personalized chronic disease care in home–hospital environments. IEEE Trans Inf Technol Biomed 2012;16(4):572–85.
[4] Moxey A, Robertson J, Newby D, Hains I, Williamson M, Pearson S-A. Computerized clinical decision support for prescribing: provision does not guarantee uptake. J Am Med Inform Assoc 2010;17(1):25–33.
[5] Kim D, Seo Y, Youn C-H. Prediction of atrial fibrillation onsets by multiple formulas based on long term variation of heart rate variability. In: IEEE int. conf. e-health networking, application & services. 2008. p. 5482–5.
[6] Lyman J, Cohn W, Bloomrosen M, Detmer D. Clinical decision support: progress and opportunities. J Am Med Inform Assoc 2010;17(5):487–92.
[7] March R. Delivering on the promise of personalized healthcare. Personalized Med 2010;7(3):327–37.
[8] Bell M, Legendre P. Multicharacter chronological clustering in a sequence of fossil sticklebacks. Syst Zool 1987;36(1):52–61.
[9] Legendre P, Dallot S, Legendre L. Succession of species within a community: chronological clustering, with applications to marine and freshwater zooplankton. Am Nat 1985;125(2):257–88.
[10] Hazemi F, Youn C, Al-Rubeaan K. Grid-based interactive diabetes system. In: IEEE int. conf. healthcare informatics, imaging and systems biology. 2011. p. 258–63.
[11] Jeong S, Youn C-H, Kim Y-W. A chronic disease identification scheme using radar chart method for personalized healthcare system. In: KIPS int. conf. ubi info tech. and appl. 2013. p. 13–20.
[12] Kahn R, Buse J, Ferrannini E, Stern M. The metabolic syndrome: time for a critical appraisal. Diabetes Care 2005;28(9):2289–304.
[13] Peng C. A hybrid cloud approach for sharing health information in chronic disease self-management. MS Thesis. Blekinge Institute of Technology; 2013.