A Study on Diagnostic Decision Support System for Metabolic Syndrome Care in Cloud Integrated Clinic Environments

정 상 진 (鄭 相 真 Jeong, Sangjin)  
정보통신공학과  
Department of Information and Communications Engineering  
KAIST  
2014
A Study on Diagnostic Decision Support System for Metabolic Syndrome Care in Cloud Integrated Clinic Environments
A Study on Diagnostic Decision Support System for Metabolic Syndrome Care in Cloud Integrated Clinic Environments

Advisor : Professor Youn, Chan-Hyun

by

Jeong, Sangjin

Department of Information and Communications Engineering

KAIST

A thesis submitted to the faculty of KAIST in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Information and Communications Engineering. The study was conducted in accordance with Code of Research Ethics\(^1\)

2013. 11. 13

Approved by

Professor Youn, Chan-Hyun

[Advisor]

\(^1\) Declaration of Ethical Conduct in Research: I, as a graduate student of KAIST, hereby declare that I have not committed any acts that may damage the credibility of my research. These include, but are not limited to: falsification, thesis written by someone else, distortion of research findings or plagiarism. I affirm that my thesis contains honest conclusions based on my own careful research under the guidance of my thesis advisor.
클라우드 통합형 헬스케어 환경에서 대사중후군 관리를 위한 의료의사결정시스템에 관한 연구

장 상 진

위 논문은 한국과학기술원 박사학위논문으로 학위논문심사위원회에서 심사 통과하였음.

2013년 11월 13일

심사위원장 윤 찬 현 (인)
심사위원 김 탁 곤 (인)
심사위원 박 홍 석 (인)
심사위원 최 준 귤 (인)
심사위원 함 진 호 (인)
ABSTRACT

Facing the increasing demands and challenges in the area of chronic disease care, a number of studies on the healthcare system which can, whenever and wherever, extract and process patient data, have been conducted. Chronic diseases are the long-term diseases and require periodic monitoring, multidimensional quantitative analysis, and the classification of patients’ diagnostic information. Among the chronic diseases, metabolic syndrome (MS) that refers to a clustering of specific cardiovascular disease risk factors whose underlying pathology is thought to be related to insulin resistance is one of the major chronic diseases in many countries including Korea, because of its relationship with the incidence of cardiovascular disease (CVD) and type II diabetes mellitus (T2DM). Different from acute disease, chronic disease such as MS requires long term care and its temporal information plays an important role to manage the status of disease. Health monitoring in out-of-hospital conditions, especially in the home environment, has drawn the attention of healthcare researchers and developers for a long time, because patients having chronic disease such as MS typically spend most time at home environment. In response to those increasing requirements, there have been a number of previous studies regarding MS, but most of them were focused on investigating the relationship between MS risk factors and the incidence of other chronic diseases such as CVD, coronary heart disease (CHD), and T2DM. Up to our knowledge, there has been no previous literature about quantifying the risk of MS, which is essential for predicting the incidence of MS in the future. To achieve this objective, it is imperative to overcome the well-known limitations of MS diagnostic definitions. There has been much effort to establish diagnostic criteria for MS, but it is known that current diagnostic criteria of MS have the following weaknesses such as no consideration for different importance among risk factors, thresholds-based binary style diagnosis, difficulty in estimating the risk of MS for non-MS subjects, and difficulty in managing the temporal change of the status of MS risk factors. This dissertation proposes a novel MS risk quantification model, which resolves the weaknesses of MS diagnosis methods and shows the validity of the model using extensive clinical evaluation using a large number of sample health examination data. We also present a temporal progress model of MS risk and cloud-based healthcare system architecture to effectively care MS patients by considering the characteristics of MS and chronic diseases.

In this dissertation, we first propose a risk quantification model for MS, which is based on areal similarity degree analysis between weighted radar charts consisting of MS diagnostic criteria and examination results of MS risk factors of a subject. The clinical effectiveness of the proposed model is evaluated using data
of a large number of subjects obtained from the third Korea National Health and Nutrition Examination Survey (KNHANES III). The evaluation results show that the proposed model can quantify the risk of MS and effectively identify a group of subjects who can be classified into a potential risk group for having MS in the future. Using the proposed model, we can early identify potential MS patients and monitor the temporal change of the patients’ statuses.

Then, we propose a patient-specific chronic disease care system using a chronological clustering method to analyze the temporal progress of chronic disease. We also present a patient tier classification method based on the sensitivity level for accepting controlling the changes of patient’s disease status. The proposed system provides personalized chronic disease care services according to the classified patient tiers. Through these new technologies, we can design new application services, such as analyzing long-term trends of patient’s disease status, a knowledge-based decision support tool for cardiovascular disease, and a system which predicts mitochondria-level metabolic disorder. The clinical effectiveness of the proposed model is also evaluated using KNHANES III data. The evaluation results show that the proposed model can be used to analyze the temporal progress of chronic disease, especially, the risk of MS. By using the temporal model, we can effectively manage patients with MS or patients at risk of developing it.

Finally, we propose a new healthcare system architecture that integrates the at-home and at-hospital environment to effectively care MS patients. The system supports customizability and dynamic functionality update in a personalized healthcare system by using cloud-based at-home and at-hospital environments integration. The proposed system architecture provides MS risk management functionality based on both MS risk quantification model and temporal progress model. The service broker module within the system supports dynamic provisioning and configuration of personalized at-home healthcare system in cloud environments. We also present the prototype implementation of the personalized healthcare system in home-hospital cloud environments.

The primary contribution and applicability to healthcare and biomedical engineering field is that it proposed an innovative method to identify potential patients with having high risk of MS in advance so that physicians and patients could proactively manage health status and reduce time and medical expenditure for caring the MS. To achieve the objective, we developed a novel MS risk quantification model based on areal similarity degree (ASD) analysis, a temporal progress model based on chronological clustering methodology, and an integrated healthcare architecture for home-hospital integrated healthcare system for MS care.

The rest of this dissertation is organized as follows. Chapter 2 investigates the characteristics of chronic disease care cycles and the prevalence of major chronic diseases. Among the diseases, we further investigate the concept and diagnostic criteria of MS and present the state-of-arts on the healthcare system for effective chronic disease care. Based on the results, we identify the major problems of the diagnostic criteria for MS and the healthcare systems for MS patients caring. Then, we propose healthcare service flows integrating home and hospital environments. In Chapter 3, we present an MS risk quantification model to identify potential MS subjects with high risk by using the novel ASD analysis, where the diagnostic criteria for MS are ambiguous, and show the clinical effectiveness in terms of applicability for preventing the incidence of CVD and
T2DM through managing the status of risk factors of the identified potential MS subjects. In Chapter 4, we propose a temporal progress model of MS risk and patients’ disease states. The effectiveness of the temporal model is also evaluated using clinical examination results of patients. Chapter 5 proposes cloud-based healthcare system architecture to effectively care MS patients by considering the characteristics of MS and chronic diseases, particularly out-of-hospital management. The prototype implementation is also presented. This dissertation is concluded and describes the limitations and further work in the Chapter 6.

Keywords: cloud, healthcare, chronic disease, metabolic syndrome, clinical decision support system, health informatics
Table of Contents

Abstract .................................................................i
Table of Contents ......................................................iv
List of Tables .........................................................vii
List of Figures .........................................................ix

Chapter 1. Introduction ........................................ 1
1.1 Overview ..........................................................1
1.1.1 Acronyms and Abbreviations ...............................3
1.2 Motivation ..........................................................5
1.3 Contributions ....................................................9
1.4 Dissertation Outline ...........................................11

Chapter 2. Integrated Healthcare Environments for Chronic Disease Patients Management ........................................ 12
2.1 Introduction ......................................................12
2.2 Backgrounds .....................................................14
2.2.1 Characteristics of Chronic Disease Care .....................14
2.2.2 Current Diagnostic Definitions of Metabolic Syndrome ....22
2.2.3 State-of-Art on Clinical Decision Support System ..........26
2.3 Problem Description ............................................34
2.3.1 Problem Description on MS Diagnostic Criteria .............34
2.3.2 Problem Description on Conventional Healthcare Environments ........................................ 40
2.4 Proposed Service Scenarios for Chronic Disease Care in Home-Hospital Environments ........................................ 53
2.4.1 Home-Hospital Integrated Healthcare Service Scenario ........53
2.4.2 Cloud-based Personalized Healthcare Service Scenario ....55
2.5 Predicted Cost Model for Integrated Healthcare Service ..........57
2.6 Summary ..........................................................62
Chapter 3. Diagnostic Decision Method for Metabolic Syndrome Based on Risk Quantification Model

3.1 Introduction ................................................................................................................. 63

3.1.1 Problem Description on Current Diagnostic Decision Method for Metabolic Syndrome .................................................................................................................. 64

3.2 Related Work on Radar Chart Model ............................................................................. 66

3.3 Novel MS Risk Quantification Model ............................................................................. 68

3.3.1 Construction of Weighted Radar Chart for Describing the Status of MS Risk Factors .................................................................................................................. 68

3.3.2 MS Risk Modeling Using Areal Similarity Degree ....................................................... 72

3.3.3 Discussion of Weighted Radar Chart Model .............................................................. 79

3.4 Evaluation and Discussion ............................................................................................ 83

3.4.1 Characteristics of Study Subjects .............................................................................. 83

3.4.2 Evaluation .................................................................................................................. 85

3.4.3 Discussion .................................................................................................................. 95

3.5 Summary ...................................................................................................................... 121

Chapter 4. Temporal Chronic Disease Progress Model for Clinical Decision Support

4.1 Introduction .................................................................................................................... 122

4.2 Problem Description ..................................................................................................... 124

4.3 Model for Temporal Progress of MS Using Chronological Clustering ....................... 125

4.3.1 Chronological Clustering Method for Identifying Temporal Variation of MS .......... 125

4.3.2 Evaluation and Discussion ....................................................................................... 130

4.3.3 ASD Value Estimation Based on Multiple Regressions ............................................ 133

4.3.4 Summary .................................................................................................................. 138

4.4 MS Status Tier Classification for Patients Using Control Range of Chronological Clustering .................................................................................................................. 139

4.4.1 Determination of Control Range Thresholds for Patient Tier Classification based on Chronological Clustering ............................................................................. 139

4.4.2 Discussion ............................................................................................................... 147

4.4.3 Summary ................................................................................................................. 148

4.5 Summary ...................................................................................................................... 149
Table of Tables

2.1 MS and age-adjusted risk for outcomes for Framingham Offspring at 8-year follow-up [47] ................................................................. 23
2.2 Diagnostic definitions of metabolic syndrome ................................................. 25
2.3 Characteristics summary of investigated health information systems ................. 33
2.4 Summary of problems for MS diagnostic criteria .............................................. 36
2.5 Prevalence and risk for CVD, CHD, and T2DM associated with specific combinations of MS risk factors relative to individuals without that combination [47] .............. 37
2.6 MS risk factors’ examination results of sample subjects .................................. 38
2.7 MS diagnosis criteria for Korean males ............................................................... 38
2.8 Regular examination results of MS risk factors .................................................... 39
2.9 Annual number of hospital visits of total Korean population ............................. 41
2.10 Characteristics of integrated healthcare system for chronic disease .................. 50
2.11 States of integrated healthcare system in home-hospital environment for the Markov process ........................................................................... 58
3.1 Pairwise comparison matrix .............................................................................. 69
3.2 Values of random index (R.I.) [23] ..................................................................... 70
3.3 Computed weights for each indicator and allocation of angle in radar chart .......... 71
3.4 Sample examination results of type 2 diabetes patient (male) ............................ 71
3.5 Thresholds of five risk factors of MS for Korean subjects ................................. 78
3.6 Characteristics of study subjects ....................................................................... 84
3.7 Calculation of weights for MS risk factors among study subjects ...................... 85
3.8 Detailed sub-cases of each sub-group belongs to male subjects ....................... 88
3.9 Examination results for MS risk factors for case 2 subjects ............................... 94
3.10 Changes in the number of Case 2, Case 4, and ASD thresholds ....................... 96
3.11 Changes of the weights of MS risk factors and ASD thresholds ....................... 99
3.12 Weights of MS risk factors for comparison ....................................................... 101
3.13 Examination results of MS risk factors for middle-aged male subjects having ASD value 1 and duplicated SM RF ......................................................... 108
3.14 Examination results of MS risk factors for middle-aged male subjects having ASD value 0.934 and duplicated SM RF ......................................................... 110
3.15 Examination results of MS risk factors for young female subjects having ASD

3.16 Characteristics of sampling subjects

3.17 MS probability table for young male subgroup

3.18 MS probability table for middle-aged and old male subgroups

3.19 MS probability table for young, middle-aged, and old female subgroups

3.20 Possibility of a subject has MS given that the ASD value exceeds the corresponding thresholds for young, middle-aged, and old female subgroups

4.1 A sample patient’s medical examination results for metabolic syndrome risk factors

4.2 Calculation of chronological distance of a sample patient

4.3 Medical examination results of Subject 1 for retrospective clinical analysis

(Young male)

4.4 Medical examination results of Subject 2 for retrospective clinical analysis

(Middle-aged male)

4.5 Summary of ASD estimation accuracy comparison

4.6 Patient’s TIER classification based on $d_{t(2)}$ thresholds

4.7 Detailed sub-cases of each subject-group [56]

4.8 Criteria for determining thresholds of patients MS status changes

4.9 ASD thresholds for male subjects sub-groups

4.10 ASD thresholds for female subjects sub-groups

5.1 Three service levels for chronic disease patients
Table of Figures

1.1 Major challenges in chronic disease diagnosis and care flow ................................. 8
2.1 An example of chronic disease care cycles ................................................................. 15
2.2 Incidences of obesity for males and females .............................................................. 16
2.3 Incidences of hypertension for males and females ...................................................... 17
2.4 Incidences of hypercholesterolemia for males and females ........................................ 18
2.5 Incidences of low HDL-cholesterol for males and females .......................................... 19
2.6 Incidences of hypertriglyceridemia for males and females .......................................... 20
2.7 Incidences of T2DM for males and females ................................................................. 21
2.8 Screenshot of Lifelines2 main interface [103] ............................................................... 27
2.9 Screenshot of Knave-II system [105] ........................................................................... 29
2.10 Screenshot of TimeLine system [106] ........................................................................ 30
2.11 Screenshot of AsbruView showing temporal trends of patient [107] ......................... 31
2.12 Integrated overall view of Sakura viewer and Fuji viewer [110] ............................... 32
2.13 Typical process for periodic hospital visits of MS patients ....................................... 40
2.14 Sources of medical information at home environment .............................................. 44
2.15 Sources of medical information at hospital environment ......................................... 46
2.16 An example screenshot from health information system at hospital ....................... 46
2.17 An example of clinical report for patient ................................................................... 48
2.18 Various types of examination results .......................................................................... 49
2.19 Implementation examples of personalized disease care service on mobile devices
    (patient view) .................................................................................................................. 52
2.20 Implementation examples of personalized disease care service on mobile devices
    (physician view) ............................................................................................................. 52
2.21 Service flow for personalized chronic disease care for the home-hospital ................ 52
2.22 Service scenario for the personalized chronic disease management in home-hospital
    clouds .............................................................................................................................. 56
3.1 Decomposition of risk factor structure for MS .......................................................... 69
3.2 Weighted radar chart for status of MS risk factors ..................................................... 72
3.3 A partial radar chart consisting of two polygons ....................................................... 73
3.4 Two weighted radar charts constructed by examination results of a subject and
3.5 Sample weighted radar charts

3.6 Exchange the positions of indicator 1 and 2

3.7 Exchange the positions of indicator 1 and 3

3.8 Calculation of ASD for two weighted radar charts

3.9 Regression analysis of proposed ASD-based model and the number of MS risk factors for male subjects sub-groups

3.10 Regression analysis of proposed ASD-based model and the number of MS risk factors for female subjects sub-groups

3.11 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for male sub-groups: young male sub-group. Interval of ASD range=0.01

3.12 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for male sub-groups: middle-aged male sub-group. Interval of ASD range=0.01

3.13 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for male sub-groups: old male sub-group. Interval of ASD range=0.01

3.14 Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for young male sub-group

3.15 Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for middle-aged male sub-group

3.16 Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for old male sub-group

3.17 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for female sub-groups: young sub-group. Interval of ASD range=0.01

3.18 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for female sub-groups: middle-aged sub-group. Interval of ASD range=0.01

3.19 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for female sub-groups: old sub-group. Interval of ASD range=0.01

3.20 Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for young female sub-group

3.21 Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for middle-aged female sub-group

3.22 Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD
value for old female sub-group ........................................... 93

3.23 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for young male sub-group of KNHANES IV-2008: risk factor weights and ASD threshold of KNHANES III-2005 (a), updated risk factor weights and ASD threshold using KNHANES IV-2008 (b). Interval of ASD range=0.01 ...................... 98

3.24 Analysis of the numbers of MS subjects and non-MS subjects over ASD value for middle-aged female sub-group of KNHANES IV-2008: risk factor weights and ASD threshold of KNHANES III-2005 (a), updated risk factor weights and ASD threshold using KNHANES IV-2008 (b). Interval of ASD range=0.01 ...................... 98

3.25 Procedures for utilizing the ASD method to other populations .................. 100

3.26 Differences of the number of Case 2 between CASE2_{EWR} and CASE2_{ASD} over ASD value ................................................................. 102

3.27 Ratio of total number of Case 2 subjects of CASE2_{EWR} to CASE2_{ASD} ........... 103

3.28 Prototype implementation of MS risk management application ................ 104

3.29 Temporal changes of ASD value during regular health examinations .......... 105

3.30 Weighted radar chart showing detailed information of MS risk factors ........ 105

3.31 Distribution of the occurrence of same ASD values for middle-aged male sub-group ................................................................................. 106

3.32 Ancillary indicator to distinguish subjects having the same ASD value ........ 107

3.33 Distribution of the occurrences of SM_{rf} for the middle-aged male subjects with ASD value 1 ........................................................................... 108

3.34 Distribution of the occurrences of SM_{rf} for the middle-aged male subjects with ASD value 0.934 ................................................................. 109

3.35 Distribution of the occurrence of same ASD values for young female sub-group .... 111

3.36 Distribution of the occurrences of SM_{rf} for the young female subjects with ASD value 0.333 ............................................................................. 111

3.37 Changes of ASD values of sampling subjects of middle-aged male subjects over time ......................................................................................... 115

3.38 Changes of ASD values of sampling subjects of middle-aged female subjects over time ......................................................................................... 115

3.39 Bayesian network model for MS risk diagnosis using ASD thresholds ........ 117

4.1 A partial radar chart consisting of two polygons. The figure has been adopted from
4.2 Two weighted radar charts constructed from two data sets A and B. The figure has been adopted from [56] and modified ................................................................. 127
4.3 Linear regression of ASD values over time ....................................................... 130
4.4 Linear regression of ASD values of Subject 1 ................................................... 132
4.5 Linear regression of ASD values of Subject 2 ................................................... 135
4.6 Application of linear regressions of MS risk factors to future ASD estimation ...... 137
4.7 Linear regression results for MS risk factors of Subject 1 ................................. 137
4.8 Linear regression results for MS risk factors of Subject 2 ................................. 138
4.9 Sensitivity for temporal change of MS status according to different values of δ ..... 140
4.10 Distributions of frequency counts of young-adult male subjects over ASD values 143
4.11 Distributions of frequency counts of middle-aged male subjects over ASD values 143
4.12 Distributions of frequency counts of old-aged male subjects over ASD values ...... 144
4.13 ASD thresholds for TIER classification of male subjects ............................... 144
4.14 Distributions of frequency counts of young-adult female subjects over ASD values 145
4.15 Distributions of frequency counts of middle-aged female subjects over ASD values 146
4.16 Distributions of frequency counts of old-aged female over ASD values .......... 146
4.17 ASD thresholds for TIER classification of female patients ........................... 147
5.1 Architecture of the personalized healthcare system for chronic disease in home-hospital cloud environments ................................................................. 155
5.2 Architecture of the integrated healthcare system ............................................. 157
5.3 Operational procedure of the MS risk prediction service ............................... 160
5.4 Operational procedure of the temporal progress care service for the MS .......... 162
5.5 Health risk quantification process ................................................................... 163
5.6 Informant-driven mental stress factor identification service ........................... 164
5.7 Operational procedure for the mental stress factor identification service .......... 165
5.8 Operational procedure for the e-Re number estimation-based heart disease detection service ...................................................................................... 166
5.9 Operational procedure for the mitochondria-level metabolic disorder prediction service ...................................................................................... 167
5.10 Service broker module for personalized healthcare system in cloud environment ····· 169
5.11 Mobile application for personalized chronic disease care service ······························ 171
5.12 Prototype implementation of a patient’s home healthcare devices ······························ 172
5.13 Temporal changes of ASD value during regular health examinations ······························ 173
Chapter 1. Introduction

This chapter introduces the motivation and major contributions of the study presented in this dissertation. It describes a high-level overview of the key concepts related to the problems addressed in the dissertation, followed by the fundamental motivations behind this dissertation.

1.1 Overview

Facing the increasing demands and challenges in the area of chronic disease care, a number of studies on the healthcare system which can, whenever and wherever, extract and process patient data, have been conducted. Chronic diseases are the long-term diseases and require periodic monitoring, multidimensional quantitative analysis, and the classification of patients’ diagnostic information. According to the national survey of Korea, the prevalence of chronic disease for adults over 20 years-old is almost 30% in 2005. Further, medical expenditure for 11 major chronic diseases was 16 trillion Korean Won (KRW) i.e., 1.3% of gross domestic production (GDP) of Korea in 2010. For the US, more than 75% of annual medical expenditure, i.e., 2 trillion US Dollar (USD) was spent on medical care of chronic disease in 2010 [20].

One of the major chronic diseases in many countries including Korea is metabolic syndrome (MS) that refers to a clustering of specific cardiovascular disease risk factors whose underlying pathology is thought to be related to insulin resistance. The risk factors include insulin resistance, obesity, dyslipidemia, and hypertension and it is known to increase the risk for cardiovascular disease (CVD) and type II diabetes mellitus (T2DM). Since MS helps to identify individuals at high risk for both cardiovascular disease and T2DM, it has become one of the major issues of public healthcare in many countries, especially developed countries. There has been much effort to establish diagnostic criteria for MS, but it has been well known that the current diagnostic criteria of MS have weaknesses such as binary decision based on diagnostic criteria, equal weight among risk factors, and difficulty in estimating the temporal progress of the risk factors. To resolve these weaknesses, it is essential to develop a method for quantifying the risk of MS, but which has not previously appeared in the MS literature. However, the development of methods for progress management of metabolic syndrome has not been widely investigated.

Different from acute disease, chronic disease such MS requires long term care and its temporal in-
formation plays an important role to manage the status of disease. Thus, a patient having MS should visit the hospital periodically, which generates highly frequent medical data and causes huge volume of medical data to be accumulated. In response to this challenge, many studies have reported various technical u-health service systems in patient-care monitoring utilizing sensor networks and medical services recently. Particularly, the development of u-health technologies such as mobile computing using dynamic software adaptation techniques [65] or new networking technologies has seen the important elements of chronic conditions based on sensors become a primary issue. In other words, through mobile devices, subcompact sensors, and wireless networks, a health examination is executed and can transmit in real time a patient’s physical data to a medical center. Therefore, these systems enable out-of-hospital health monitoring.

Health monitoring in out-of-hospital conditions, especially in the home environment, has drawn the attention of healthcare researchers and developers for a long time, because patients having chronic disease such as MS typically spend most time at home environment. Sensors and measurement devices for health monitoring at home have limited capability for accurate analyses of measured physiological data, and the analysis processes are performed by predefined methods in the devices. Furthermore, it is not easy for a patient to understand the analysis results fully due to a lack of medical knowledge. Therefore, it is important work to allow chronic patients to manage their own conditions these days, and healthcare systems are required to assist patients’ self-management of their chronic condition by delivering more precise information and suggesting suitable disease management methods. Patients need a long treatment period with continuous monitoring care. Their condition sometimes may change or worsen unexpectedly. Thus, it is important from the patients’ point of view to provide a customizable self-management capability for patients having MS. However, most of these studies have been based on independent healthcare systems that operate in either the home or the hospital. Therefore, it is required to develop new healthcare system integrating two different healthcare systems, i.e., hospital and home healthcare system seamlessly.
1.1.1 Acronyms and Abbreviations

This dissertation uses the following acronyms and abbreviations:

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AACE</td>
<td>American Association of Clinical Endocrinologists</td>
</tr>
<tr>
<td>AHA/NHLBI</td>
<td>American Heart Association - National Heart, Lung, and Blood Institute</td>
</tr>
<tr>
<td>AHP</td>
<td>Analytic Hierarchy Process</td>
</tr>
<tr>
<td>ASD</td>
<td>Areal Similarity Degree</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>BP</td>
<td>Blood Pressure</td>
</tr>
<tr>
<td>CDA</td>
<td>Clinical Document Architecture</td>
</tr>
<tr>
<td>CDSS</td>
<td>Clinical Decision Support System</td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary Heart Disease</td>
</tr>
<tr>
<td>CI</td>
<td>Consistency Index</td>
</tr>
<tr>
<td>CR</td>
<td>Consistency Ratio</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular Disease</td>
</tr>
<tr>
<td>DBP</td>
<td>Diastolic Blood Pressure</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiography</td>
</tr>
<tr>
<td>EGIR</td>
<td>European Group for the Study of Insulin Resistance</td>
</tr>
<tr>
<td>EHR</td>
<td>Electronic Health Record</td>
</tr>
<tr>
<td>FG</td>
<td>Fasting Glucose</td>
</tr>
<tr>
<td>HDL-C</td>
<td>High Density Lipoprotein - Cholesterol</td>
</tr>
<tr>
<td>HIMS</td>
<td>Health Information Management System</td>
</tr>
<tr>
<td>HIRA</td>
<td>Health Insurance Review &amp; Assessment Service</td>
</tr>
<tr>
<td>IDF</td>
<td>International Diabetes Federation</td>
</tr>
<tr>
<td>IGT</td>
<td>Impaired Glucose Tolerance</td>
</tr>
<tr>
<td>KNHANES III</td>
<td>the third Korea National Health and Nutrition Examination Survey</td>
</tr>
<tr>
<td>KNHANES IV</td>
<td>the fourth Korea National Health and Nutrition Examination Survey</td>
</tr>
</tbody>
</table>
mDNA: mitochondrial DeoxyriboNucleic Acid
MS: Metabolic Syndrome
NCEP ATP III: National Cholesterol Education Program - Third Adult Treatment Panel
PACS: Picture Archiving and Communication System
PAR: Population-Attributable Risk
PHR: Personal Health Record
PSCM: Patient Status Classification Method
RCA: Radar Chart Approach
RI: Random Index
RIS: Radiology Information System
RR: Relative Risk
SBP: Systolic Blood Pressure
SLA: Service Level Agreement
SM: Sensitivity Metric
SMOP: Surface Measure of Overall Performance
SOAP: Simple Object Access Protocol
SRI: Stress Response Inventory
T2DM: Type II Diabetes Mellitus
TG: Triglycerides
UDDI: Universal Description, Discovery and Integration
WAS: Web Application Server
WC: Waist Circumference
WHO: World Health Organization
WSDL: Web Service Definition Language
XML: eXtensible Markup Language
1.2 Motivation

MS has become a major public healthcare issue in many countries because of its relationship with major chronic disease such as T2DM and CVD. There has been much effort to establish diagnostic criteria for MS among several expert groups such as WHO, EGIR, NCEP ATP III, AACE, AHA/NHLBI, and IDF. All the diagnostic criteria include obesity, insulin resistance, dyslipidemia, and hypertension as the critical risk factors of MS. However, the detailed definition and thresholds for each critical risk factor are different, so the existence of multiple definitions for MS has caused confusion, and it has proved difficult to make direct comparisons between studies in which different definitions were used to identify the syndrome or the data from different countries. Among the MS diagnosis definitions, NCEP ATP III definition is popularly used definition. NCEP ATP III definition comprises the following criteria.

Three or more of the following five risk factors:

- Fasting Glucose (FG): ≥ 110 mg/dl but non-diabetic
- Blood Pressure (BP): ≥ 130/85 mmHg
- Triglycerides (TG): ≥ 150 mg/dl
- High Density Lipoprotein - Cholesterol (HDL-C): Male < 40mg/dl
  Female < 50mg/dl
- Obesity (Waist Circumference (WC)): Male > 102 cm
  Female > 88 cm

According to the diagnostic definition, a patient is diagnosed to have MS if the patient’s health examination results exceed the designated thresholds defined above. However, since MS is a chronic disease, which shows temporal progress and leads to major complications over time, its expression depends heavily on age and exposure to an unhealthy lifestyle. Diagnosis of MS should not be managed as an acute infection, and a binary decision based on a yes or no approach is not valid in chronic degenerative disorders such as MS. MS incorporates clinical entities in which insulin resistance plays a major pathological role. This concept results in a group of subjects with a spectrum of long-term risk for having the final outcomes. The risk depends on the characteristics of the population being examined. Future adaptations to the current definitions of MS should
take into account ethnic variability. Selection of diagnostic thresholds should be supported by the risk for developing major complications, especially if treatment can prevent these outcomes. The following case is a typical example to show the weakness of current MS diagnostic criteria. A 55-year old man who had a body mass index (BMI) of 27 kg/m\(^2\), a waist circumference of 98 cm, HDL-cholesterol of 36 mg/dl, triglycerides of 180 mg/dl, fasting glucose of 90 mg/dl, post-challenge 2 hours glucose of 180 mg/dl, fasting insulin of 32 \(\mu\)U/ml, and a blood pressure of 120/80 mmHg is not determined to have MS by either the NCEP ATP III or the WHO definition, despite the presence of insulin resistance, impaired glucose tolerance, hypertriglyceridemia, low HDL-cholesterol, and overweight [2][4].

Therefore, in order to increase the effectiveness and usefulness for predicting the prevalence of MS using the current diagnostic definition, it is essential to resolve the following limitations of current MS diagnostic criteria.

**No consideration for different importance among risk factors**

It is well known that the risk factors comprising MS have different effect to the incidence of major complications such as CVD, T2DM, and so on. However, the current diagnostic definitions do not consider the different importance among risk factors.

**Thresholds-based binary style diagnosis**

The fact that there are cut points for the various risk factors implies that values above the specified thresholds are associated with excess risk, yet the rationale for the specific cut points, as opposed to higher or lower values, has never been delineated.

**Difficulty in estimating the risk of MS for non-MS subjects**

An individual is diagnosed as having MS if the individual has three or more risk factors that exceed the defined thresholds, for example. However, if a person has two risk factors, the person is not diagnosed as MS patients even though the person has a risk of developing MS in the future. The current diagnostic criteria
do not provide a capability for identifying those potential MS patients who has high risk for having MS in the future.

**Difficulty in managing the temporal change of the status of MS risk factors**

Since patients with chronic diseases including MS typically stay outside of hospitals for the majority of chronic disease care cycles except routine hospital visits, providing application for personal devices can help patients with self-management. However, it is difficult for a patient to know whether the patient’s health status has been improved or not over time, because the current definitions provide the number of risk factors only. For example, let us assume that an individual has two MS risk factors FG and WC. Then, the health examination results of the risk factors will be changed over time, but the current definitions cannot provide any information about the temporal changes of the risk factors.

Therefore, the major problems observed above should be resolved in order to effectively care MS patients. Further, it is necessary to develop a healthcare system that can help with MS patients perform self-management at home environment during their home-care period. Also, to maximize the caring effectiveness of MS patients, the healthcare system should integrate at-home and at-hospital systems and provide collaborative environment between patients and physicians.

Fig. 1.1 summarizes the research motivation of this dissertation. The key objective of this dissertation is to identify potential patients with having high risk of chronic diseases in advance so that physicians and patients can proactively manage health status and reduce time and medical expenditure for caring the chronic diseases. It is noted that this dissertation focuses on MS among many chronic diseases because of its predictability and relationship with the prevalence of CVD and T2DM. When an individual visits hospital for health examination, a physician typically performs disease diagnosis, operations, therapy, and prescription. Among the physician’s roles, this research focuses on disease diagnosis, particularly chronic disease, because other roles require clinical knowledge and performing medical operations by non-physicians is restricted. During the diagnosis process, the physician identifies the individual’s chronic diseases by using health examination results and healthcare system in the hospital. If the individual already If the individual already If the individual
already had chronic diseases and visited the hospital after out-of-hospital caring, the physician checks the out-of-hospital caring status and determine the individual’s health status change. After the diagnosis procedures, the individual is determined as normal or chronic diseases patient. However, in the current chronic disease diagnosis method, particularly MS, it is difficult to determine if the individual has potential risk for having MS. Thus, it is necessary to develop a method for identifying normal individuals with high risk for having chronic diseases in advance. When the individual is diagnosed as having chronic diseases, the individual returns to home and performs self-management at home environment. After the self-management period, the individual visits the hospital and checks the temporal changes of the individual’s disease status. 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. To achieve the objective, we develop a novel MS risk quantification model based on areal similarity degree (ASD) analysis, a temporal progress model based on chronological clustering methodology, and an integrated healthcare architecture for home-hospital integrated healthcare system for MS care.

Figure 1.1: Major challenges in chronic disease diagnosis and care flow.
1.3 Contributions

This dissertation contributes to developing an innovative MS risk quantification model, which has not been appeared from previous MS literature and to developing a new healthcare system architecture, which helps with MS patients and physicians managing the disease status effectively. The contributions are as follows:

1) The primary contribution of the dissertation is to develop a novel risk quantification model for MS, which is based on ASD analysis between weighted radar charts consisting of MS diagnostic criteria and examination results of MS risk factors of a subject. The clinical effectiveness of the proposed ASD model was evaluated using data of a large number of subjects obtained from the nationwide health examination. The evaluation results showed that the proposed ASD model can quantify the risk of MS and effectively identify a group of subjects who can be classified into a potential risk group for having MS in the future. We further elaborate the proposed MS risk quantification model to resolve various theoretical and clinical issues on the proposed ASD model. We have showed the clinical effectiveness of the proposed ASD model by comparing with the conventional MS diagnosis method and presented clinical benefits and applicability to other populations. We have additionally developed an ancillary method for complementing the weaknesses of the proposed ASD model and clinically distinguishing individuals having the same risk by our proposed MS risk quantification model. We also extend the ASD model by applying Bayesian network model to determining the reference ASD threshold values for identifying potential MS patients.

2) Secondly, among the known weaknesses, to resolve the difficulty in managing the temporal changes of MS status, we propose a 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 of ASD between consecutive MS risk factors examination results. We evaluated the proposed temporal change model using a sample patient’s examination results. 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. 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.

3) In the third contribution, we proposed a new healthcare system architecture that integrates the at-home and at-hospital environment to effectively care MS patients. The system supports customizability and dynamic functionality update in a personalized healthcare system by using cloud-based at-home and at-hospital environments integration. The proposed system architecture provides MS risk management functionality based on both MS risk quantification model and temporal progress model. The service broker module within the system supports dynamic provisioning and configuration of personalized at-home healthcare system in cloud environments. Then, we presented Markov process-based cost model to predict costs based upon healthcare service scenario of home-hospital integrated environments.

4) Finally, we also present the prototype implementation of the personalized healthcare system in home-hospital cloud environments. For easy operation, we have implemented the smartphone user interface and connected the interface with the personalized healthcare system in the testbed. We expect the proposed personalized healthcare system and user interface to be applicable to provide cost effective and personalized chronic disease care.

The primary contribution and applicability to healthcare and biomedical engineering field of this dissertation is that it proposed an innovative method to identify potential patients with having high risk of MS in advance so that physicians and patients could proactively manage health status and reduce time and medical expenditure for caring the MS. To achieve the objective, we developed a novel MS risk quantification model based on ASD analysis, a temporal progress model based on chronological clustering methodology, and an integrated healthcare architecture for home-hospital integrated healthcare system for MS care.
1.4 Dissertation Outline

The rest of this dissertation is organized as follows. Chapter 2 investigates the characteristics of chronic disease care cycles and the prevalence of major chronic diseases. Among the diseases, we further investigate the concept and diagnostic criteria of MS and present the state-of-arts regarding the healthcare system for caring chronic disease effectively. Based on the survey results, we identify the major problems regarding the diagnostic criteria for MS and the healthcare systems for MS patients caring. Then, we propose healthcare service flows integrating home and hospital environments. In Chapter 3, we present an MS risk quantification model to identify potential MS subjects with high risk by using the novel ASD analysis, where the definitions and diagnostic criteria for MS are ambiguous, and show the clinical effectiveness in terms of applicability for preventing the incidence of MS, CVD, and T2DM through managing the status of risk factors of the identified potential MS subjects. In Chapter 4, we propose a temporal progress model of MS risk and patients’ disease status. The effectiveness of the temporal model is also evaluated using clinical examination results of patients. Chapter 5 proposes a cloud-based new healthcare system architecture to effectively care MS patients by considering the characteristics of MS and chronic diseases, particularly out-of-hospital management. The healthcare system is evaluated in terms of predicted cost for MS caring along with prototype implementation of the healthcare system for MS patients. This dissertation is concluded and describes the limitations and further work in the Chapter 6.
Chapter 2. Integrated Healthcare Environments for Chronic Disease Patients Management

This chapter presents backgrounds of this dissertation including characteristics of chronic disease care, current diagnostic definitions of MS, and state-of-art on healthcare systems and clinical decision support systems. By reviewing the relevant literatures, we identify major problems on MS diagnostic criteria and conventional healthcare environments. Then, we propose service scenarios for effective chronic disease care in home-hospital environments. The scenarios are comprised of two sub-scenarios, home-hospital integrated healthcare service scenario and cloud-based personalized healthcare service scenario, respectively. Then, we present Markov model-based predicted cost model for the proposed chronic disease care service scenarios.

2.1 Introduction

Facing the increasing demands and challenges in the area of chronic disease care, various studies on the healthcare system which can, whenever and wherever, extract and process patient data, have been conducted. Chronic diseases are the long-term diseases and require periodic monitoring, multidimensional quantitative analysis, and the classification of patients’ diagnostic information. According to the national survey of Korea, the prevalence of chronic disease for adults over 20 years-old is almost 30% in 2005. Further, medical expenditure for 11 major chronic diseases was 16 trillion Korean Won (KRW) i.e., 1.3% of gross domestic production (GDP) of Korea in 2010. For the US, more than 75% of annual medical expenditure, i.e., 2 trillion US Dollar (USD) was spent on medical care of chronic disease in 2010 [20]. One of the major chronic diseases in many countries including Korea is metabolic syndrome (MS). MS refers to a clustering of specific cardiovascular disease risk factors whose underlying pathology is thought to be related to insulin resistance. The risk factors include insulin resistance, obesity, dyslipidemia, and hypertension and it is known to increase the risk for cardiovascular disease (CVD) and type II diabetes mellitus (T2DM). Since MS helps to identify individuals at high risk for both cardiovascular disease and T2DM, it has become one of the major issues of public healthcare in many countries. There has been much effort to establish diagnostic criteria for metabolic syndrome, but current diagnostic criteria of metabolic syndrome have weaknesses such as binary decision based on diagnostic criteria, equal weight among risk factors, and difficulty in estimating the temporal progress of
the risk factors.

As a response to these demands, healthcare systems have been evolved to more effectively manage patients with chronic diseases. A healthcare system for chronic diseases is characterized as an at-hospital and at-home service according to a targeted environment. Both services basically aim to provide patients with accurate diagnoses of disease by monitoring a variety of physical states with a number of monitoring methods, but there are differences between home and hospital environments, and the different characteristics should be considered in order to provide more accurate diagnoses for patients, especially, patients having chronic diseases. Furthermore, the development of an integrated and personalized healthcare system is emerging as an important issue in the modern healthcare industry. There have been several studies that have sought to provide integrated healthcare services in home and hospital environments. However, as the concept of cloud computing has become mature, there is an increasing demand to utilize a large data archive of clinical data records, decision support systems, and the event-based notification and monitoring system to achieve more accurate diagnosis and disease care. Also, metabolic syndrome has become a major issue of public healthcare in many countries. There are efforts to develop the metabolic syndrome risk quantification and prediction model and to integrate them to personalized healthcare system, so as to predict the risk of having metabolic syndrome in the future.

In this chapter, we first review the previous literature on the chronic diseases, metabolic syndrome, and healthcare systems for managing chronic disease patients in considering the diverse environments. Then, we investigate the major problems and limitations not yet resolved by the current studies, particularly diagnostic criteria for MS and architecture of healthcare system for effective chronic disease care. We propose two healthcare service scenarios for effective caring of chronic disease by considering the characteristics of chronic disease caring environments. The proposed service scenarios are evaluated by using Markov model-based cost prediction.
2.2 Backgrounds

This subchapter presents the characteristics of caring chronic diseases and the incidence of major chronic diseases. Then, we investigate the concept, diagnostic criteria, and clinical importance of MS and the state-of-arts of healthcare system for caring chronic diseases effectively.

2.2.1 Characteristics of Chronic Disease Care

Chronic diseases are increasingly an important concern in e-healthcare systems throughout the world. For example, it is forecasted that clinical expenses for chronic diseases in the U.S. will be 80% of total medical costs and that more than 150 million people may suffer from chronic diseases in 2020 [64]. To effectively care the increasing incidence of chronic disease patients, there is a major trend in modern healthcare industry. It is the changing focus of healthcare. Within a decade, the “baby boomers” will begin reaching the age of retirement, and the percentage of Americans age 65 or older will jump from 13% to 20% of the total population. Due to the aging population, healthcare delivery will increasingly move its focus from acute, episodic care to treatment of long-term, chronic conditions. By the year 2020, 157 million Americans will be treated for a chronic condition; those treatments will represent 80% of total U.S. healthcare expenditures. This patient population - aging with chronic needs, increasingly sophisticated, and accustomed to the instant gratification made possible by the digital age - will have correspondingly high expectations of their healthcare system. The second trend is the computerization and interconnection of medical devices and systems. Computerization of medical devices is a natural consequence of the need to rapidly acquire, process, and present an ever-increasing quantity and variety of healthcare information. Interconnecting disparate medical devices can yield a more direct, accurate, and rapid exchange of healthcare information. Combined, these trends will increasingly move us away from the traditional offline hospitals and physician’s offices and move us toward their virtual successors. Ultimately, diagnosis and treatment will take place at an earlier stage, with less impact on the patient’s lifestyle, and with more effect as healthcare delivery occurs in the home, office, school, and other non-intrusive community settings [64].

Different from acute disease, chronic disease requires long term care and its temporal information plays an important role to manage the status of disease. Thus, a patient having chronic disease should visit the hospital periodically, which generates highly frequent medical data and causes huge volume of medical data to
be accumulated. Fig. 2.1 shows typical chronic care cycles for a chronic disease patient who is required to have routine visits every time interval $T$. Except visiting the hospital, the patients stay outside of the healthcare setting for the majority of the chronic care cycle. Over time, a chronic disease can become degenerative, and a patient’s home healthcare activities may become more intensive. During time $T_1$ and $T_2$, the patient routinely visits the hospital for periodic examination. However, at time $T_2$, the patient’s disease status becomes degenerative, so the physician requests the patient to visit the hospital more frequently. At time $T_4$, the patient’s disease status becomes stable, so the physician extends the routine visit interval. This scenario shows typical example of chronic disease care, so as shown in the figure, patient spends most time at home and performs self-management of the patient’s own chronic disease. The patient generates a lot of medical examination data during home healthcare, which cause large computational overhead on the health information systems in the hospital. Therefore, it is important to develop effective self-management systems at the patient’s home environment and to develop new health information system architecture for integrating different environments.

![Figure 2.1: An example of chronic disease care cycles.](image)

To investigate the incidences of chronic diseases in populations, we have performed the analyses of incidences of major chronic disease in the Korean population. The analyses are based on data obtained from the third Korea National Health and Nutrition Examination Survey (KNHANES III) among non-institutionalized civilians in the Republic of Korea, which was conducted by the Korean Ministry of Health and Welfare in 2005. This survey was a nationwide representative study using a stratified, multistage probability sampling design for the selection of household units. The survey consisted of the following 4 components:
the Health Interview Survey, the Health Behavior Survey, the Health Examination Survey, and the Nutrition Survey [27]. A total of 34,145 individuals from these sampling frames were included in the health interview survey; among them, 25,161 subjects aged over 20 years were identified as potential participants in our study.

We analyzed the incidences of major chronic diseases, such as obesity, hypertension, hypercholesterolemia, low HDL-cholesterol, hypertriglyceridemia, and T2DM. Fig. 2.2 shows the incidences of obesity. The diagnostic criterion for obesity is the body mass index (BMI) ≥ 25.0 kg/m². The average BMI of male subjects who are greater than or equal to 20 years old is 24.0 kg/m². For female subjects, the BMI is 23.3 kg/m². For male subjects, the forties and the fifties show peak average BMI, whereas for female subjects, the average BMI increases proportional to ages until the sixties. Then, the average BMI decreases in seventies. The incidences of obesity are 35.2% and 28.3% for male and female subjects, respectively. Male subjects show high incidence of obesity at forties and fifties, whereas female subjects show high incidence at fifties and sixties. The incidences of underweight are 3.6% and 5.7% for male and female subjects, respectively. Male subjects show the peaks incidence at over seventies, whereas female subjects show at twenties.

Figure 2.2: Incidences of obesity for males and females.
Fig. 2.3 shows the incidences of hypertension. The diagnostic criterion for hypertension is systolic blood pressure (SBP) \( \geq 140 \) mmHg, diastolic blood pressure (DBP) \( \geq 90 \) mmHg, or in treatment. The average SBP of male subjects who are greater than or equal to 10 years old is 119.0 mmHg. For female subjects, the average SBP is 113.3 mmHg. Both males and females show positive correlation between the average SBP and ages. Particularly, females show stronger correlation between the average SBP and ages than males. The average DBP of male subjects who are greater than or equal to 10 years old is 78.4 mmHg. For female subjects, the average DBP is 73.2 mmHg. The average values of DBP for male subjects increase until fifties, then the average values of DBP decrease. For females, the average values of DBP decrease after sixties. The incidences of hypertension are 30.2% and 25.6%, respectively. The incidences of male subjects increase until sixties, then the incidences decrease. For females, the incidences are proportional to the ages. It is noted that the incidences are over 50% for males in sixties and females over 60 years old. Self-recognition rates of hypertension are 47.8% for males and 65.9% for females. In treatment rates are 39.2% and 60.0% for males and females, respectively. These results indicate that females perform better self-management than males.

Figure 2.3: Incidences of hypertension for males and females.
Fig. 2.4 shows the incidences of hypercholesterolemia. The diagnostic criterion for hypercholesterolemia is total cholesterol $\geq$ 240 mg/dl or in treatment. The average total cholesterol of male subjects who are greater than or equal to 10 years old is 178.6 mg/dl. For female subjects, the average total cholesterol is 180.0 mg/dl. For male subjects, the average value increases until forties, then it decreases. For female subjects, the average value increases until sixties and the average value of females over fifty years old is higher than males’. The incidences of hypercholesterolemia are 7.5% and 8.8% for males and females, respectively. Both males and females show increasing incidence of hypercholesterolemia until sixties. Then, the incidences decrease.

![Graph showing incidences of hypercholesterolemia for males and females.](image)

Figure 2.4: Incidences of hypercholesterolemia for males and females.
Fig. 2.5 shows the incidences of low HDL-cholesterol. The diagnostic criterion for low HDL-cholesterol is HDL-cholesterol < 40 mg/dl. The average HDL-cholesterol of male subjects who are greater than or equal to 10 years old is 42.4 mg/dl. For female subjects, the average HDL-cholesterol is 47.4 mg/dl. For male subjects, the average value does not show correlation with ages, however female subjects show decreasing trend over twenties. The incidences of low HDL-cholesterol are 47.9% and 26.0% for males and females, respectively. The incidence of male subjects is almost twice of female subjects’. This trend is kept until the fifties.

Figure 2.5: Incidences of low HDL-cholesterol for males and females.
Fig. 2.6 shows the incidences of hypertriglyceridemia. The diagnostic criterion for hypertriglyceridemia is triglycerides ≥ 200 mg/dl. The average value of triglycerides for males who are greater than or equal to 10 years old is 145.1 mg/dl. For females, the average value of triglycerides is 108.6 mg/dl. Both males and females show positive correlation between the average value and ages. The average values of the forties show the greatest difference between males and females, 183.3 mg/dl versus 112.4 mg/dl. The incidences of hypertriglyceridemia are 23.7% and 10.7% for males and females, respectively. The incidences of males show peak values at the forties and the fifties. For females, the peak values are appeared at the fifties and the sixties.

![Figure 2.6: Incidences of hypertriglyceridemia for males and females.](image-url)
Fig. 2.7 shows the incidences of T2DM. The diagnostic criterion for T2DM is fasting glucose $\geq 126$ mg/dl or in treatment. The average fasting glucose of male subjects who are greater than or equal to 10 years old is 94.3 mg/dl. For female subjects, the average fasting glucose is 91.6 mg/dl. Both males and females show positive correlation between the average fasting glucose until the fifties. However, males show decreasing trend after sixties, whereas females keep increasing trend until seventies. The incidence of T2DM for males is 9.0%, whereas that of females is 7.2%. In accordance with the average fasting glucose, females show increasing trend until seventies. Self-recognition rates of T2DM are 58.08% for males and 66.5% for females. In treatment rates are 48.6% and 59.3% for males and females, respectively. These results indicate that females self-recognize their health status. However, males show better glucose control rates than females, 26.2% versus 19.4%, respectively.

As we investigated in this subchapter, the major chronic diseases show high prevalence, so it is important to effectively care the patients having chronic diseases. It is further required to diagnose potential patients in advance in order to perform proactive therapy and care.
2.2.2 Current Diagnostic Definitions of Metabolic Syndrome

There are a number of literatures that introduce the concept and clinical importance of the metabolic syndrome (MS). Among them, in this subchapter, we refer to the explanation presented in [111]. The MS integrates, in a single diagnosis, the clinical manifestations of insulin resistance and obesity that lead to increased cardiovascular morbidity and precede T2DM [112]. Clustering those risk factors has advantages over the analysis of the same entities by themselves. In the previous studies, the relative risk for having coronary heart disease (CHD) is significantly greater in cases with the MS compared to the risk associated with each component alone. This effect is confirmed in cases with normal or impaired glucose tolerance [113] and in patients with T2DM [114]. Botnia et al. performed a 6.9-year follow-up study of 4483 first-degree relatives of patients with T2DM, the relative risk for having CHD was greater in patients with the MS (relative risk (RR) = 2.96) compared to that found for obesity (RR = 1.44), dyslipidemia (RR = 1.73), hypertension (RR = 1.57), microalbuminuria and insulin resistance by themselves. The same is true for incident diabetes [115]. Although impaired glucose tolerance (IGT) is the major contributor to the risk for T2DM in the future, the addition of at least two other elements of the metabolic syndrome (e.g., high triglycerides, low HDL cholesterol and hypertension) further increased the precision of the estimation of future diabetes. Thus, the MS concept is a strong conglomerate of risk factors that are useful for identifying subjects at risk for future CHD and/or diabetes. In addition, the concept of the MS concept provides an integrative view of the pathophysiology of the disease. The risk factors of MS are markers of abnormalities in insulin-regulated metabolic pathways referred schematically as traits. For each trait, there exist multiple studies showing the relationship of insulin resistance with its pathophysiology. In several instances, insulin resistance acts in more than one way to induce the trait. For example, in the dyslipidemia of the metabolic syndrome, insulin resistance causes both increased synthesis and decreased clearance of the liver-derived triglyceride-rich lipoproteins, both mechanisms resulting in abnormally high triglyceride concentrations [116]. Table 2.1 shows the 8-year follow-up study results of the clinical relationship between MS and other chronic disease. According to the study, MS is associated with an increased risk for CHD, CVD, and T2DM in both males and females. Also, MS accounted for up to one third of CVD in male and approximately half of new T2DM cases over 8 years of follow-up study [47].
Table 2.1: MS and age-adjusted risk for outcomes for Framingham Offspring at 8-year follow-up [47].

<table>
<thead>
<tr>
<th>Event</th>
<th>No. of events/Nonevents, MS Absent</th>
<th>No. of events/Nonevents, MS Present</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td>CVD</td>
<td>53/1081 (4.7%)</td>
<td>63/352 (15.2%)</td>
</tr>
<tr>
<td>Hard CHD</td>
<td>23/1111 (2.0%)</td>
<td>25/390 (6.0%)</td>
</tr>
<tr>
<td>Total CHD</td>
<td>38/1096 (3.4%)</td>
<td>40/375 (9.6%)</td>
</tr>
<tr>
<td>T2DM</td>
<td>28/1106 (2.5%)</td>
<td>71/344 (17.1%)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>CVD</td>
<td>37/1442 (2.5%)</td>
<td>21/274 (7.1%)</td>
</tr>
<tr>
<td>Hard CHD</td>
<td>8/1471 (0.5%)</td>
<td>5/290 (1.7%)</td>
</tr>
<tr>
<td>Total CHD</td>
<td>21/1458 (1.4%)</td>
<td>8/287 (2.7%)</td>
</tr>
<tr>
<td>T2DM</td>
<td>33/1446 (2.2%)</td>
<td>46/249 (15.6%)</td>
</tr>
</tbody>
</table>

However, the pathophysiology of the traits is not completely explained by the existence of insulin resistance. For example, the higher than expected prevalence of primary dyslipidemias (e.g., familial combined hyperlipidemia or familial hypertriglyceridemia) found in cases with the metabolic syndrome cannot be explained solely by insulin resistance. Based on this, the World Health Organization (WHO) selected the name “metabolic syndrome” instead of the insulin resistance syndrome [117]. Therefore, the MS is a valuable tool in medical education; the concept provides a broad overview of the complex mechanisms by which insulin resistance causes long-term complications [111].

However, the complex nature of insulin resistance has resulted in controversy. Insulin resistance is defined as a decreased glucose lowering response to insulin and is usually associated with an abnormal insulin secretion pattern that results in increased plasma concentration of mature and immature forms of the hormone [118]. Additional complexity comes from the multiple and diverse metabolic pathways regulated by insulin and the selectivity of the defect to some tissues. The hormone is the major determinant of protein and lipid
synthesis in the liver and muscle. It is the main negative regulator of lipolysis in the adipose tissue. In addition, it regulates the synthesis of diverse compounds such as coagulation factors, sex hormones and apolipoproteins. Insulin coordinates global responses such as the lipoprotein sources during fasting or feeding, tissue growth or the endothelial function. In the metabolic syndrome, the severity of the defect varies between tissues and metabolic pathways. The liver, muscle, endothelial cells and adipose tissue are the main affected organs [119]. Even in the same tissue, the defect is limited to only some insulin-regulated processes. For example, in the muscle, insulin resistance is associated with decreased glucose transport and storage, but no changes occur in constituent protein synthesis. The coexistence of increased insulin concentrations makes the pathophysiology of the syndrome even more complex. Some of the manifestations are caused by a decreased insulin action, i.e., hyperglycemia; in contrast, others are explained by excessive amounts of insulin, i.e., acanthosis nigricans. Finally, clinical manifestations may vary among subjects with the same degree of insulin resistance. Two subjects with the same insulin sensitivity index may have widely different triglyceride concentration or HDL cholesterol levels. Thus, neither insulin resistance nor the plasma insulin concentration is the only determinant of the expression of the disease. The intricate nature of insulin resistance leads to multiple symptoms and various clinical pictures. Hence, the metabolic syndrome concept integrates this diversity in a single entity with well-known outcomes [111].

Apart from investigating the clinical relationship among traits of MS, there has been much effort to establish diagnostic criteria for MS among several expert groups such as the WHO, the European Group for the Study of Insulin Resistance (EGIR), the National Cholesterol Education Program - Third Adult Treatment Panel (NCEP ATP III), the American Association of Clinical Endocrinologists (AACE), and American Heart Association - National Heart, Lung, and Blood Institute (AHA/NHLBI), and the International Diabetes Federation (IDF). Table 2.2 shows the current clinical definitions of MS. All the diagnostic criteria include obesity, insulin resistance, dyslipidemia, and hypertension as the critical risk factors of MS. However, the detailed definition and thresholds for each critical risk factor are different, so the existence of multiple definitions for MS has caused confusion, and it has proved difficult to make direct comparisons between studies in which different definitions were used to identify the syndrome or the data from different countries. Therefore, many studies are being performed to develop a consensus about the definition and thresholds for diagnostic criteria so as to acquire more accurate diagnosis [1].

- 24 -
Table 2.2: Diagnostic definitions of metabolic syndrome.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose intolerance, IGT or diabetes and/or insulin resistance together with two or more of the following:</td>
<td>Glucose intolerance, IGT or diabetes and/or insulin resistance together with two or more of the following:</td>
<td>Insulin resistance together with two of the following:</td>
<td>Three or more of the following five risk factors:</td>
<td>Obesity and two or more of the following four risk factors:</td>
<td>Three or more of the following five risk factors:</td>
</tr>
<tr>
<td>FG</td>
<td>≥ 110 mg/dl but non-diabetic</td>
<td>≥ 110 mg/dl but non-diabetic</td>
<td>≥ 100 mg/dl or treatment</td>
<td>≥ 100 mg/dl or treatment</td>
<td>≥ 100 mg/dl or treatment</td>
</tr>
<tr>
<td>BP</td>
<td>≥ 140/90 mmHg or treatment</td>
<td>≥ 140/90 mmHg or treatment</td>
<td>≥ 130/85 mmHg or treatment</td>
<td>≥ 130/85 mmHg or treatment</td>
<td>≥ 130/85 mmHg or treatment</td>
</tr>
<tr>
<td>TG</td>
<td>≥ 150 mg/dl and/or treatment</td>
<td>≥ 150 mg/dl</td>
<td>≥ 150 mg/dl</td>
<td>≥ 150 mg/dl</td>
<td>≥ 150 mg/dl</td>
</tr>
<tr>
<td>HDL-C</td>
<td>Male: &lt; 35 mg/dl or treatment</td>
<td>&lt; 39 mg/dl or treatment</td>
<td>Male: &lt; 40 mg/dl or treatment</td>
<td>Male: &lt; 40 mg/dl or treatment</td>
<td>Male: &lt;40mg/dl or treatment</td>
</tr>
<tr>
<td></td>
<td>Female: &lt; 39 mg/dl</td>
<td>Female: &lt; 40 mg/dl</td>
<td>Female: &lt; 50 mg/dl</td>
<td>Female: &lt; 50 mg/dl</td>
<td>Female: &lt;50mg/dl or treatment</td>
</tr>
<tr>
<td>Obesity</td>
<td>Male: WHR &gt; 0.9, Female: WHR &gt; 0.85 and/or BMI &gt; 30 kg/m²</td>
<td>Male: WC ≥ 94 cm</td>
<td>Male: WC &gt; 102 cm</td>
<td>Male: WC &gt; 102 cm</td>
<td>Male: WC &gt; 102 cm</td>
</tr>
<tr>
<td></td>
<td>Female: WC ≥ 80 cm</td>
<td>Female: WC &gt; 88 cm</td>
<td>Female: WC &gt; 88 cm</td>
<td>Female: WC &gt; 88 cm</td>
<td>Female: WC &gt; 88 cm</td>
</tr>
<tr>
<td>Microalbuminuria</td>
<td>Urinary albumin excretion rate ≥ 20μg/min or albumin/creatinine ratio ≥ 30mg/g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IGT, impaired glucose tolerance; FG, fasting glucose; BP, blood pressure; TG, triglycerides; HDL-C, high-density lipoprotein cholesterol; WHR, waist-hip ratio; WC, waist circumference; BMI, body mass index
2.2.3 State-of-Art on Clinical Decision Support System

In response to challenge about chronic disease, many studies have reported various technical e-health service systems in patient-care monitoring utilizing sensor networks and medical services recently. Particularly, the development of e-health technologies such as mobile computing using dynamic software adaptation techniques or new networking technologies has seen the important elements of chronic conditions based on sensors become a primary issue. In other words, through mobile devices, subcompact sensors, and wireless networks, a health examination is executed and can transmit in real time a patient’s physical data to a medical center. In this subchapter, we investigate the state-of-art systems related to health information systems especially from the viewpoint of effectively managing temporal health information of a large number of chronic disease patients. Our objective is to survey the key health information systems whether the systems provide effective management capability of temporal disease status. The management capability includes representation of medical histories, visual data query and aggregation, generation of temporal abstractions and visualization of treatment plans.

A. Lifelines

LifeLines uses a timeline visualization technique to represent personal histories, medical records and other types on biographical data. In LifeLines, horizontal bars are used to indicate the temporal timeline and events are displayed on the timeline. Similar events such as similar prescriptions are organized into facets, which can be expanded and collapsed to provide increasing or decreasing level of detail. Color notations and line thickness are used to indicate the importance and relationship of events. To handle regions with high data density, LifeLines provides zooming functionality allowing users to compress and stretch the time scale at any location. Additional content such as x-ray images, ultrasonic sound, etc. can be added in a linked fashion. Authors apply LifeLines in the analysis of complex patient medical records to visualize temporal relationships between treatments, consultations, disorders, prescriptions, hospitalizations and other events [102].
B. Lifelines2

LifeLines2 is an extended version of LifeLines, which allows the user to analyze records from multiple patients at a time. The system focuses on providing comparative visualization of records by means of aligning, filtering and sorting operations. By arranging patient records according to target reference event, physicians can easily point out concurrent or neighboring events. To provide interactive use, ranking and filtering operations are supported and customizing the set of records to suit according to a physician’s requirements is possible. It is known that the system is particularly suitable for observational research, where researchers analyze data from different studies in order to better understand health problems or study the effect of treatments, and in finding patients for clinical trials. Fig. 2.8 shows the screenshot of Lifeline2 health information system [103].

Figure 2.8: Screenshot of Lifelines2 main interface [103].
C. CLEF

CLEF health information system is proposed by Hallet et al. to browse medical histories, which integrates visual navigation tools and conveniently generated textual summaries. Even though the graphical interface is used to interactively investigate the temporal trends, textual descriptions are also useful for representing complex temporal information. Within the CLEF system, the patient’s medical history is represented as a network of semantically and temporally organized events, which serves as an input for visualization and natural language generation components. The visual navigator depicts a high level overview of a patient’s medical history by plotting events along three parallel timelines, corresponding to diagnoses, treatments and investigations. In addition to zooming time scale and detail-on-demand functionality, the navigator provides interactive visualization of semantic relationships between events. Having different features from the LifeLines interface, the navigator also allows the user to visualize numerical data by plotting results of measurements on separate line charts. Natural language generation is used for two purposes: 1) to create customized textual reports for printing or exchange purposes and 2) as a support tool for the visual navigator, to enable better description of complex events and relationships between them [104].

D. KNAVE-II

KNAVE-II is an interface enabling knowledge-based visualization and interactive exploration of time-oriented data at different levels of temporal abstractions. Users can navigate through the links of a semantic network while simultaneously navigating visually through multiple degrees of temporal abstraction of the dataset under observation. The evaluation results have shown that users of KNAVE-II were able to perform queries both faster and more accurately than with other standard tools. Figure 2.9 displays the captured screenshot of KNAVE-II system [105].
E. TimeLine

The TimeLine system is a problem-centric temporal visualization system for patient records. The contents of the electronic health records are integrated, reorganized, and displayed within the user interface on the timeline. It is similar to Lifelines in the way that the different elements of the health records such as imaging, reports, lab tests, etc. are grouped along the vertical axis. However, the major difference from Lifeline2 system is that the TimeLine system uses an XML data representation format to handle data from distributed, heterogeneous medical databases. Data elements that are displayed in the user interface are classified based on a knowledge base that guides both data inclusion rules and the visualization metaphors used to render the data. Figure 2.10 shows a sample screenshot of Timeline health information system [106].
Figure 2.10: Screenshot of TimeLine system [106].

F. ASBRUVIEW

AsbruView is a visualization and user interface on top of Asbru language-based health information representation and treatment system to show the structured time-oriented plans between physicians and patients. AsbruView represents hierarchical and temporal relationships between treatment plans using a 3D visualization perspective. Plans are aligned along the time axis and can be stacked on top of each other and laid out in different ways. To provide simple interface, AsbruView provides well-known real world objects-based graphic user interface. Also a 2D view is available which focuses on temporal aspects of plans in greater detail. To depict uncertainty of future events, AsbruView extends the timeline by using time annotation glyphs. Figure 2.11 depicts a sample screenshot showing temporal trends of patient status, which is captured from AsbruView user interface [107-108].
G. SAKURA-Viewer

The Sakura Viewer focuses the view of consolidated information. It visualizes chronological order data through the diverse user interfaces including the patient status progress view. The Sakura viewer shows order history from two viewpoints simultaneously. This eliminates any semantic redundancies in the data, such as a repeatedly issued order item. The Sakura viewer is intended to reduce the volume of data to be displayed, by consolidating data that occur repeatedly in the order history and by eliminating any redundancy. The Sakura viewer is a highly effective tool, as: 1) it visualizes both the semantic viewpoint and the temporal viewpoint of patient records simultaneously; 2) it promotes awareness of contextual information among the daily data; and 3) it implements patient-centric data entry methods. This viewer contributes to decrease the physician’s workload in an order entry system. [109].

H. FUJI-Viewer

The Fuji Viewer focuses the flow of periodic information. It can be integrated with Sakura viewer system and used to visualize order data and order plan data through the periodic view. This viewer shows differences between the plan history and the order history. This viewer builds for only test order-related data. Because the test order and the test plan are correlated, it is necessary to manage a long-term test order history and test plan history concurrently. This viewer supports the abstract view for test order related data. This
viewer can promote the user’s awareness of test item’s periodicity in the test order related data. This viewer supports users’ understanding of the repeating patterns and trends of test items. Figure 2.12 shows the integrated overall view of Sakura viewer and Fuji viewer [110].

![Figure 2.12: Integrated overall view of Sakura viewer and Fuji viewer [110.]](image)

Most of the investigated health information systems are directed at clinicians and clinical practice, although they are not always developed in close relation to them. Table 2.3 shows the characteristics summary of investigated health information systems to give an overview of intended users for each system. Also, their proposed goals and tasks are listed. From the physician point of view, a number of tasks and goals can be defined for each system. Some are very specific and tend to care for niche usages, while others provide more general visualization methods that can be used for general purposes.
### Table 2.3: Characteristics summary of investigated health information systems.

<table>
<thead>
<tr>
<th>System</th>
<th>User, Goals, and Tasks</th>
</tr>
</thead>
</table>
| **Lifelines** | - Physician  
|           | - Patient care  
|           | - Use electronic health record in temporal time-based view  |
| **Lifelines2** | - Clinical researcher  
|             | - Research  
|             | - Compare patterns of events, detecting trends  |
| **CLEF**   | - Physician, biomedical researcher  
|           | - Patient care  
|           | - Visualize timelines, use natural language processing to extract complex temporal data, aggregate numerical data  |
| **KNAVE-II** | - Physician  
|           | - Patient care  
|           | - Generation and exploration of context sensitive abstraction of temporal data  |
| **TimeLine** | - Physician  
|            | - Patient care  
|            | - Use electronic health record content in temporal time-based view with additional filters on data based on natural language processing technology  |
| **Asbru View** | - Physician  
|               | - Patient care  
|               | - Medical therapy planning and execution  |
| **Sakura viewer** | - Physician  
|                       | - Patient care  
|                       | - Manage order history to identify redundancy  |
| **Fuji viewer** | - Physician  
|                | - Patient care  
|                | - Manage periodical history of order and plan. Can be integrated with Sakura Viewer  |
2.3 Problem Description

In this subchapter, we investigate the major problems of the conventional MS diagnostic criteria and healthcare systems.

2.3.1 Problem Description on MS Diagnostic Criteria

MS has become a major public healthcare issue in many countries. MS refers to a cluster of specific CVD risk factors whose underlying pathology is thought to be related to insulin resistance. Those risk factors include insulin resistance, obesity, dyslipidemia, and hypertension and they are thought to increase the risk for CVD and T2DM. Therefore, the importance of MS is that it helps to identify individuals at high risk for both CVD and T2DM [1][2].

There has been much effort to establish diagnostic criteria for MS among several expert groups such as WHO, EGIR, NCEP ATP III, AACE, AHA/NHLBI, and IDF. All the diagnostic criteria include obesity, insulin resistance, dyslipidemia, and hypertension as the critical risk factors of MS. However, the detailed definition and thresholds for each critical risk factor are different, so the existence of multiple definitions for MS has caused confusion, and it has proved difficult to make direct comparisons between studies in which different definitions were used to identify the syndrome or the data from different countries. Table 2.1 summarizes the current definitions of MS [51]. Therefore, many studies are being performed to develop a consensus about the definition and thresholds for diagnostic criteria so as to acquire more accurate diagnosis [1].

There have been many studies to investigate the risk factor structure of MS, the prevalence of MS, the impact of the risk factors in the development of MS, and the relationship between MS and other diseases. Some of these studies are found in [3-10]. Also, there is another approach to investigate the mechanism of MS from the microscopic perspective at the cell and mitochondria level [11-13]. Those previous studies have mostly aimed to analyze the effects of individual risk factors on MS, but due to the clustering characteristics of MS, it is necessary to take a holistic approach for a more accurate understanding of MS. Multivariate analysis has shown that the individual risk factors comprising the syndrome each carry different odds ratios for predicting the prevalence of CHD, the incidence of CHD, and CVD mortality. Peter et al. showed that MS is associated with an increased risk for CHD, CVD, and T2DM in both males and females. They found that MS accounted for up to one third of CVD in male and approximately half of new T2DM cases over 8 years of fol-
McNeill et al. also showed that individuals without diabetes or CVD, but with MS, were at increased risk for long-term cardiovascular outcomes. Thus, the identification of individuals with MS may provide opportunities to intervene earlier in the development of shared disease pathways that predispose individuals to both CVD and T2DM [48]. In addition to hyperglycemia, low high-density lipoprotein (HDL) cholesterol and hypertension usually indicate a significantly greater risk compared with the presence of obesity or high triglycerides [14][15]. The studies also illustrate another likely shortcoming of the current approach to diagnosing MS. All the NCEP ATP III, WHO, and AHA/NHLBI definitions weigh each risk component equally, yet it is clear that some risk factors included in the definitions have greater CVD predictive value than others. This fact is highlighted in other algorithms used to predict CVD risk using regression coefficients to assign different weights to risk factors [16][17], and it is apparent from studies that examined the risk of CVD in persons with one or two components of MS versus three or more [18].

According to these studies, MS should be considered a progressive process that leads to major complications over time. Its expression depends heavily on age and exposure to an unhealthy lifestyle. Diagnosis of MS should not be managed as an acute infection, and a binary decision based on a yes or no approach is not valid in chronic degenerative disorders such as MS. MS incorporates clinical entities in which insulin resistance plays a major pathological role. This concept results in a group of subjects with a spectrum of long-term risk for having the final outcomes. The risk depends on the characteristics of the population being examined. Future adaptations to the current definitions of MS should take into account ethnic variability. Selection of diagnostic thresholds should be supported by the risk for developing major complications, especially if treatment can prevent these outcomes. The following case is a typical example to show the weakness of current MS diagnostic criteria. A 55-year old man who had a BMI of 27 kg/m², a waist circumference of 98 cm, HDL-cholesterol of 36 mg/dl, triglycerides of 180 mg/dl, fasting glucose of 90 mg/dl, post-challenge 2 hours glucose of 180 mg/dl, fasting insulin of 32 μU/ml, and a blood pressure of 120/80 mmHg is not determined to have MS by either the NCEP ATP III or the WHO definition, despite the presence of insulin resistance, impaired glucose tolerance, hypertriglyceridemia, low HDL-cholesterol, and overweight [2][4]. The weaknesses of the current diagnostic criteria of MS are summarized in Table 2.4.
Table 2.4: Summary of problems for MS diagnostic criteria.

- 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

A. No consideration for different importance of factors

The relative importance among MS risk factors for the incidence of relevant chronic disease such as CVD, CHD, and T2DM is extensively investigated in the study of Wilson et al. [47]. The authors investigated the potential effects of presence of various combinations of MS risk factors from single to pairs or triplets on the incidences of CVD, CHD, and T2DM [47]. This clause presents the summarized results of [47]. Table 2.5 summarizes the analysis results. Risk for outcomes associated with specific risk factors combinations was estimated with the group without that specific combination used as the comparator. This analysis demonstrated the heterogeneity in distribution of the various risk factor combinations that make up the aggregate MS and the associated heterogeneity in risk for outcomes. The entries in the top of Table 2.5 show the age- and sex-adjusted relative risks (RRs) for single risk factors. For example, the MS blood pressure criterion (model 1.4) was present in 48.8% of the participants and imparted an RR of 2.0 for CVD events, using a comparison group of persons without the MS blood pressure criterion. The MS risk factors that contributed most to CVD outcomes were blood pressure and HDL cholesterol, with population-attributable risk (PAR) estimates of 33% (model 1.4 in Table 2.5) and 25% (model 1.5 in Table 2.5), respectively.

Prediction of T2DM shown in models 1.1 through 1.5 with each of the MS risk factors taken individually showed similar effects, with PAR estimates of 30% to 43%, except for impaired fasting glucose, which was associated with a very large 12-fold increased risk for incident T2DM and a PAR of 62%. Different combinations of 2 and 3 MS risk factors, shown at the bottom of Table 2.5, provided estimates for the outcomes, but the prevalence of some combinations was small. Analyses based on a larger number of persons at risk or that combined groups were more informative. For example, models 3.1 through 3.6 (Table 2.5), which included impaired fasting glucose were all highly related to the development of T2DM and had RRs for vascular disease outcomes that were much lower. The composite model 3a (Table 2.5) synthesized information for all
groups that included impaired fasting glucose and 2 additional metabolic syndrome traits. This trait grouping was present in 8.9% of individuals and was associated with a very high RR for T2DM during follow-up (RR=11.0) and less impressive RRs (2.1 to 2.5) for vascular disease events. Model 3b synthesized information for all groups, including a large waist circumference but not including impaired fasting glucose. This risk factor grouping was present in 13.2% of individuals and was associated with an elevated RR for T2DM (5.0), and risk for incident CVD was also increased approximately 2-fold [47].

Table 2.5: Prevalence and risk for CVD, CHD, and T2DM associated with specific combinations of MS risk factors relative to individuals without that combination [47].

<table>
<thead>
<tr>
<th>Model</th>
<th>MS Risk Factors</th>
<th>% of Subjects</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FG</td>
<td>TG</td>
<td>WC</td>
</tr>
<tr>
<td>1 MS Risk Factor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 MS Risk Factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>2.3</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2.6</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.8</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2.9</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.10</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>3 MS Risk Factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3.2</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>3.3</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>3.4</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>3.5</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3.6</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>3.7</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>3.8</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>3.9</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>3.10</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>3a</td>
<td>Model 3.1, 3.2, 3.3, 3.4, 3.5, or 3.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b</td>
<td>Model 3.7, 3.8, or 3.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
B. Difficulty in estimating the risk for individual not exceeding the thresholds

Since MS is diagnosed by counting the number of risk factors of which values exceed the corresponding thresholds, it is difficult to estimate the risk of an individual who has less than three risk factors exceeding the designated thresholds. Table 2.6 lists MS risk factors’ examination results of sample subjects. According to the MS diagnostic criteria shown in Table 2.7, the subjects are not diagnosed as MS. However, by examining the MS risk factors’ examination results, the subjects are diagnosed as having the following diseases or health risk. According to the examination results, it is observed that Subject 1 has hypertension due to high BP values and Subject 2 has hyperlipidemia due to elevated TG and lowered HDL-C. However, these two subjects are not diagnosed as having MS by using the conventional MS diagnosis method. Furthermore, Subject 3 does not have any risk factor of which value exceeds the designated thresholds, so the subject is generally diagnosed as having no issue on the health status. However, the examination results of risk factors are close to the corresponding thresholds, so the subject may need proactive treatment and improvement on lifestyle. But, it is difficult to identify this type of subject by using the conventional MS diagnosis method.

Table 2.6: MS risk factors’ examination results of sample subjects.

<table>
<thead>
<tr>
<th>Risk Factors</th>
<th>Subject 1 (Male)</th>
<th>Subject 2 (Male)</th>
<th>Subject 3 (Male)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waist Circumference (WC) (cm)</td>
<td>90</td>
<td>85</td>
<td>88</td>
</tr>
<tr>
<td>Triglycerides (TG) (mg/dl)</td>
<td>140</td>
<td>200</td>
<td>148</td>
</tr>
<tr>
<td>HDL-Cholesterol (HDL-C) (mg/dl)</td>
<td>41</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>Blood Pressure (BP) (mmHg)</td>
<td>140/90</td>
<td>120/80</td>
<td>125/80</td>
</tr>
<tr>
<td>Fasting Glucose (FG) (mg/dl)</td>
<td>95</td>
<td>97</td>
<td>99</td>
</tr>
</tbody>
</table>

Table 2.7: MS diagnosis criteria for Korean males.

<table>
<thead>
<tr>
<th>Risk Factors</th>
<th>Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waist Circumference (WC) (cm)</td>
<td>≥ 90 (for male), ≥ 85 (for female)</td>
</tr>
<tr>
<td>Triglycerides (TG) (mg/dl)</td>
<td>≥ 150mg/dl</td>
</tr>
<tr>
<td>HDL-Cholesterol (HDL-C) (mg/dl)</td>
<td>&lt; 40 (for male), &lt; 50 (for female)</td>
</tr>
<tr>
<td>Blood Pressure (BP) (mmHg)</td>
<td>≥ 130/85mmHg</td>
</tr>
<tr>
<td>Fasting Glucose (FG) (mg/dl)</td>
<td>≥ 100mg/dl</td>
</tr>
</tbody>
</table>
C. Difficulty in managing temporal change of disease progress

Chronic diseases such as MS require regular medical examinations in order to care the disease status, so an individual usually examines his/her health status periodically. Table 2.8 shows an example of periodic medical examination results of a sample subject. As shown in the table, the subject is not diagnosed as MS by using the conventional MS criteria, but the health status of the subject gradually becomes worse. The examination results of Month 4 indicate two risk factors, i.e. HDL-C and FG have exceeded the designated thresholds values. If the subject’s health status trend is identified earlier than Month 4, it would be possible to perform proactive care of the subject’s health status and the health status might become better at Month 4. However, the current MS diagnosis method is difficult for supporting the early identification of MS risk.

Table 2.8: Regular examination results of MS risk factors.

<table>
<thead>
<tr>
<th>Risk Factors</th>
<th>Month 1</th>
<th>Month 2</th>
<th>Month 3</th>
<th>Month 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC (cm)</td>
<td>85</td>
<td>86</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>TG (mg/dl)</td>
<td>138</td>
<td>148</td>
<td>143</td>
<td>147</td>
</tr>
<tr>
<td>HDL-C (mg/dl)</td>
<td>43</td>
<td>41</td>
<td>42</td>
<td>39</td>
</tr>
<tr>
<td>BP (mmHg)</td>
<td>120/80</td>
<td>120/80</td>
<td>115/78</td>
<td>120/80</td>
</tr>
<tr>
<td>FG (mg/dl)</td>
<td>97</td>
<td>95</td>
<td>99</td>
<td>100</td>
</tr>
</tbody>
</table>

D. Separated environment between home and hospital

Since patients having chronic diseases typically spend most of time at home environment and perform self-management, exchanging self-management results with physician during the hospital visit is crucial for caring the patients’ health status. Fig. 2.13 indicates a typical procedure for periodic hospital visits of MS patients. When a patient visits a hospital, a physician queries self-management results during at-home caring to the patient. The patient typically verbally explains at-home caring results to the physician because at-home caring results are not integrated with at-hospital healthcare system. However, during the diagnosis process, it is difficult for the physician to carefully investigate the self-management results due to limited time & cost [127]. Also, it is difficult to verbally explain the at-home caring results. To resolve this issue, clinical decision support system for MS diagnosis and temporal change analysis is required.
As we investigated in this clause, current MS diagnosis methods have several problems, which limit the effective diagnosis and caring of it. Therefore, the major problems of MS diagnosis method investigated in this clause should be resolved in order to manage the disease effectively.

2.3.2 Problem Description on Conventional Healthcare Environments

In the previous Subchapter 2.2.3, we surveyed major health information systems in terms of temporal disease progress supporting especially aimed at usability for physicians caring patients. According to our survey, all of them are focused on visualizing temporal data in a timeline, while displaying specific events from the patient data. Although directed at physicians in their daily patient care routine, they are not always developed with user feedback. Evaluation of the different tools was often based on situations outside of the clinical setting, and might not reflect reality. A more intimate dialog with physicians would benefit the creation of targeted systems addressing specific needs of the medical community. 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
A. Very limitedly available time for problem analysis

According to Health Insurance Review & Assessment Service (HIRA), it is investigated that the number of annual hospital visits for total Korean population is over 4 Billion visits in 2012. Table 2.9 shows the annual hospital visits of total Korean population [125].

Table 2.9: Annual number of hospital visits of total Korean population.

<table>
<thead>
<tr>
<th>Year</th>
<th>Hospital visits (Billion)</th>
<th>Pharmacy visits (Billion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>2.91</td>
<td>4.13</td>
</tr>
<tr>
<td>2008</td>
<td>3.40</td>
<td>4.14</td>
</tr>
<tr>
<td>2009</td>
<td>3.67</td>
<td>4.40</td>
</tr>
<tr>
<td>2010</td>
<td>3.81</td>
<td>4.46</td>
</tr>
<tr>
<td>2011</td>
<td>3.93</td>
<td>4.49</td>
</tr>
<tr>
<td>2012</td>
<td>4.15</td>
<td>4.59</td>
</tr>
<tr>
<td>Total</td>
<td>21.87</td>
<td>26.21</td>
</tr>
</tbody>
</table>

Also, it is investigated that the total number of hospitals in Korea is about 81,000 as of 2010. By using Table 2.5 and the number of hospitals, we can approximately calculate the average number of patients per hospital per day, as follows:

Let us consider the number of hospitals and the total number of hospital visits in Korea in 2010 shown above. Then, the average number of patients per hospital per day \( (A_D) \) can be calculated as:

\[
A(D)_D = \frac{3,810,000,000}{81,000 \times 365\text{ (days)}} = 128.9 .
\]  

(2.1)

If we assume that a physician works 8 hours per day, we can get average diagnostic time (min) per patient as follows:

\[
T(D)_P = \frac{60\text{ (min)} \times 8\text{ (hour)}}{A(D)_D} = 3.7 \text{ (min)} .
\]

(2.2)

If we assume that the physician works 5 days in a week, then Eq. (2.2) becomes as follows:
Therefore, the above investigation indicates that a physician has less than 3 minutes for diagnosing a patient in average.

If we consider more large sized hospital, we still get similar analysis results.

It is investigated that the average number of patients visits to big 5 Korean hospitals per day is about 10,000. Let us consider the Seoul National University Hospital that is the largest hospital in Korea. According to the hospital’s statistics, the annual number of hospital visits is 2,125,094 in 2012 and 1,312 physicians are working in the hospital [126].

According to the hospital, we can obtain the following statistics.

- The average number of hospital visits for internal medicine department per day: 2,546 patients \( (2.5) \)
- The average number of physicians working in the department for examining a patient: 21 physicians \( (2.6) \)

By using Eq. \((2.5)\) and \((2.6)\), we can calculate the average time for diagnosing a patient as follows.

\[
A(SNU)_D = \frac{2546}{21} = 121.2 .
\]

\[
T(SNU)_P = \frac{60(\text{min}) \times 8(\text{hour})}{A(SNU)_D} = 4.0 (\text{min}) .
\]

Therefore, we can confirm that there are not significant difference between Eq. \((2.4)\) and Eq. \((2.8)\).

However, as we investigated in the Subchapter 2.2.1, a patient having chronic disease visits hospital periodically. For example, typical hospital visit interval for diabetes patient with controlled status is 6 months. Thus, during the home healthcare period, the patient performs self-management, examines risk factors for diabetes such as fasting glucose, weight, blood pressure, etc. every day, and stores the examination results. When the patient visits the hospital, the patient consults the physician within a few minute and has to explain 180 medical examination results during home healthcare period (6 months). Therefore, it is very difficult for the physician to understand the patient’s self-management status during at home healthcare. For example, according to the diagnosing guidelines for chronic disease patients, it is recommended at least 10 minutes for periodic routine diagnosis a diabetes patient in USA [127].
Therefore, we can understand that the current health information systems cannot provide enough disease care service for chronic disease patients. Further, we can conclude that it is necessary to develop a new health information system that can point out important events happened during home healthcare period within very limitedly available diagnosing time.

B. Disconnected environments for patient’s chronic disease management

Currently, most healthcare information systems focus solely on patient’s management at hospital, whereas most chronic disease care services consider patient’s self-management practices at home environment. The interplay between hospital and home healthcare information systems is neglected by most healthcare systems. However, as we investigated in Subchapter 2.2.1, due to the fundamental characteristics of chronic disease, it is required to integrate two disconnected healthcare environments into a single management environment. We investigate this issue in this subchapter. In response to these situations, Chen studied to understand the use of clinical and homecare information in maintaining illness trajectories and to explore the usage patterns of personal health information in the disease management process. The findings of the study indicated that health information was received, synthesized and used by patients following repeated hospital and home healthcare cycles in the life-long process of chronic care cycles [120].

Although hospital visits consume small portion of time during a chronic disease care cycle, the information exchanged during the hospital visit is crucial for a managing patient’s health status, particularly self-management at home. Let us consider typical periodic hospital visit scenario. At the beginning of a hospital visit, the physician checks the electronic health record for the patient in the hospital health information system. Then, the physician queries to the patient about the self-management at home. The patient explains the self-management results by means of verbal communication and/or presents the self-management information at home to help the physician make clinical decision as shown in Fig. 2.14. However, due to the time limit for patient diagnosis, it is very difficult for the physician to carefully investigate the self-management results at home environment. At the end of hospital visit, the clinical information generated by the physician flows back to the patient in order to direct the next cycle of disease management at home. To do this, physicians first give verbal instructions to patients, and then leave the patients a diagnostic rationale with copies of prescriptions.

As we investigated in the previous subchapter, home monitoring information and knowledge of daily life behaviors are essential for the physician in understanding and interpreting the patient’s disease progress,
and accomplishing the diagnostic work. This is because home healthcare information consists of daily behavioral information and sentimental details that can only be described by the patient. However, the home healthcare information is not integrated with the healthcare system in the hospital, so it is difficult for the physician and patient to refer the self-management results at home environment because in most cases the usability of self-management are depend on the patient’s own efforts for recording the detailed information [120].

In other words, the depth of the information relies on how much the patient endeavors on storing the detailed information. For both the patient and the physician, healthcare information should be viewed and used as a whole. The division between clinical information, produced at hospital, and home healthcare information, derived from the patient’s daily disease management causes the at home healthcare information usually to be ignored. Instead, demarcations in health information only occurred in different cycles of chronic care, i.e., at hospital environment. Fig. 2.15 and Fig. 2.16 show examples of health information system for storing medical examination results at hospital. To utilize the patient’s home healthcare information at hospital, it is necessary to import the home healthcare information to the health information system at hospital, and vice versa [120].

Figure 2.14: Sources of medical information at home environment.
Figure 2.15: Sources of medical information at hospital environment.

Figure 2.16: An example screenshot from health information system at hospital.
Over the past several years, there has been a remarkable upsurge in activity promoting the adoption of electronic health records (EHRs). All levels of government - federal, state, regional, and local - as well as the private sector, have encouraged EHR adoption. By contrast, personal health record (PHR) systems have not received the same level of attention. While EHR system functions to serve the information needs of health care professionals, PHR systems capture health data entered by individuals and provide information related to the care of those individuals. Personal health records include tools to help individuals take a more active role in their own health. In part, PHRs represent a repository for patient data, but PHR systems can also include decision-support capabilities that can assist patients in managing chronic conditions. Most consumers and patients receive care from many health care providers, and consequently their health data are dispersed over many facilities’ paper- and EHR-based record systems. A fragmented system of storing and retrieving essential patient data impedes optimal care [121-122]. However, not only technical and business concerns of PHR-related stakeholders, but also privacy and security concerns of patients hinder the deployment of PHR [123].

Therefore, to resolve the problem above, the development of a new integrated healthcare information system for home and hospital environment is necessary. According to our investigation, it is observed that the current health information systems cannot provide enough disease care service for chronic disease patients. Further, we can conclude that it is necessary to develop a new health information system that can effectively point out important events happened during home healthcare period within diagnosing time.

C. Computational overhead due to large volume of health information data

In hospitals, each hospital maintains patient health record to achieve the following purposes.

— Patient’s health history: a record of the patient’s health status and the health services provided over time.
— Method for clinical communication and care planning among the individual health care practitioners serving the patient.
— Legal document describing the health care services provided.
— Source of data for clinical, health services, and outcome research
— Major resource for health care practitioner education.
The patient health records typically consist of the following datasets.

- Patient identification: patient identification, contact, address, etc.
- Administrative forms: medical expenses, consent to treatment, insurance information, etc.
- Clinical forms: describes patient’s status, diagnosis results, treatment
- Operative forms: includes medical report related to operation, such as anesthesia report, recovery room report, operative report, pathology report
- Obstetrical forms: includes medical reports such as antepartum record, labor and delivery record, post-partum record
- Neonatal data: includes birth history, neonatal identification, neonatal physical examination, neonatal progress note
- Nursing forms
- Ancillary forms

As we investigated, medical reports consist of many specific sub-reports, so a lot of computational workload happens to perform complex and diverse processing in time. Also, to achieve better performance, unified platform for handling medical reports is strongly required. Fig. 2.17 shows an example of clinical report. The basic report contains diagnosis results, examination results, treatment results, etc. Most of the basic data are represented as text format. However, the medical examination results include various multimedia data, a lot of computation workload is necessary to compute and analyze examination results. Fig. 2.18 shows various types of medical images and examination results.
Figure 2.17: An example of clinical report for patient.
Among the examination results, electrocardiography (ECG) signal analysis is one of typical methods to detect cardiovascular diseases, but it is known that the analysis process of ECG signal requires large computing power [87]. Similarly, other pathological data analysis requires a lot of computing resources in order to perform within limited diagnosing time. There have been many attempts to create healthcare system or platform by combining the existing medical service with IT such as sensors and communications technology. In other words, the general health clinical data and information are exchanged electronically, and patients, physicians, hospitals and laboratories can be connected through communication network to perform medical practices. Such healthcare system is enabled and supported by IT, yet the core technology will be high-speed information exchange technology and high performance computing technology in distributed environment. Cloud technology can meet these requirements by bringing heterogeneous resources together and allocating them efficiently to applications [124].
D. Lack of support for user-centric customizable healthcare system

There have been many studies regarding personalized healthcare services that aim to deliver the right treatment to the right patient at the right dose and at the right time [60]. Moreover, the development of an integrated and personalized healthcare system is becoming an important issue in the modern healthcare industry. There have been several studies that have sought to provide integrated healthcare services in home and hospital environments [20].

Table 2.10 shows the characteristics of integrated healthcare system for chronic disease in home and hospital environments. The healthcare system for chronic disease is characterized in terms of at-hospital and at-home service according to the targeted environment. Both services aim to provide patients with accurate diagnoses of diseases by monitoring a variety of physical states with a number of monitoring methods. For example, a patient can be examined using many types of medical equipment and tests at the hospital, such as X-rays, computed tomography (CT), multi-channel ECG, blood tests, magnetic resonance imaging (MRI), and others, whereas the patient may be examined with a very limited range of equipment or medical tests at home due to the price of portable medical equipment for the home, the difficulty in operating the medical equipment, or for other reasons. Therefore, extending health monitoring from the hospital to the home environment should consider different monitoring characteristics from those at the hospital [124].

Table 2.10: Characteristics of integrated healthcare system for chronic disease.

<table>
<thead>
<tr>
<th>Goal</th>
<th>At home environment</th>
<th>At hospital environment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- Personalized healthcare</td>
<td>- Efficient healthcare</td>
</tr>
<tr>
<td></td>
<td>- Provide patient customized home healthcare service</td>
<td>- Provide cost effective hospital healthcare system</td>
</tr>
<tr>
<td>Requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Easy to use</td>
<td>- Coordination with home</td>
</tr>
<tr>
<td></td>
<td>- Inter-personal communication</td>
<td>- Physician and medical staff performance</td>
</tr>
<tr>
<td></td>
<td>- Coordination with hospital</td>
<td>- Increase in hospital performance</td>
</tr>
<tr>
<td></td>
<td>- Disease information</td>
<td>- Cost effectiveness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Quality of care</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Employer and customer satisfaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Adaptability to various medical devices/disease care services</td>
</tr>
</tbody>
</table>
The main objective of at-hospital health status monitoring for patients with a chronic disease is how to provide chronic disease-specific monitoring services, as the criteria and methodology for health monitoring may vary depending on the chronic disease of the patient. For example, the measurement result of a blood pressure test may be interpreted differently based on the patient’s status, i.e., whether the patient has diabetes or not. Thus, it is important to identify the patient’s chronic disease accurately and to take disease-specific characteristics into account in at-hospital service processes.

On the other hand, one of the goals of at-home health status monitoring is to provide a simple way to monitor health status at a low cost and with low operational overheads. Assuming that two patients are monitored for the possibility of heart disease using ECG equipment at home and that Patient 1 has a history of heart attack whereas Patient 2 has no history of heart attack, in such a case, it is necessary to monitor and analyze Patient 1’s ECG signal carefully, and Patient 1 may need to use ECG equipment with complex functionalities and a high price. However, Patient 2 may use simple ECG equipment that provides a basic analysis of the ECG signal. Thus, it is necessary to provide the patient’s status based on health monitoring in order to reduce the cost and overhead of the at-home service.

Also, several issues and problems of conventional personalized healthcare system have been reported in [66]. For example, lack of customizable system for patient status and disease type means that depending on specific disease care, e.g. lung tumor, neural ecology, and etc., visualization methods should be used individually. Difficulties in discerning information from large datasets imply that in a laboratory or clinical situation, diverse types of information are used by specialists. Problems stemming from a large dataset can lead to a lack of concentration on their intentions. In particular, displaying information in the temporal domain is highly complicated. This implies that much information is not always feasible. For example, unified interface problem that means diverse types of structured information can be provided in different ways may occur.

In the clinical field, physicians utilize medical information from information systems, such as a picture archiving and communication system (PACS), a radiology information system (RIS), a health information management system (HIMS) [73], and other types with which to diagnose patients or to establish treatment plans. This information can be represented on a screen with visualization schemes for effective use. However, due to the diversity of the data and the protocols associated with clinical information, physicians have difficulties when attempting to access information from diverse sources [73]. Fig. 2.19 and Fig. 2.20 show implemen-
tation examples of patient and physician interfaces for visualizing patients’ medical information, respectively.

![Patient's Setting](image1)

![HOSPITAL](image2)

![HEALTH MANAGER](image3)

![Temporal Progress of Metabolic Syndrome Risk Analysis](image4)

(a) Patient’s interface
(b) Chronic disease care service

Figure 2.19: Implementation examples of personalized disease care service on mobile devices (patient view).

![Poincare Plot features](image5)

![Temporal Progress of Metabolic Syndrome Risk Analysis](image6)

(c) Physician’s interface
(d) Chronic disease care service

Figure 2.20: Implementation examples of personalized disease care service on mobile devices (physician view).

Therefore, it is necessary to develop a healthcare system that resolves the problems investigated in this subchapter.
2.4 Proposed Service Scenarios for Chronic Disease Care in Home-Hospital Environments

2.4.1 Home-Hospital Integrated Healthcare Service Scenario

An integrated healthcare system that enables health monitoring and disease management in the home environment has been a major research area for healthcare researchers. Integrated healthcare systems mainly focus on monitoring patients’ health status, detecting and managing potential diseases in the early stage, and managing health problems in daily life [78]. Extending health monitoring from the hospital to the home environment should not be seen as a replication of the same monitoring procedures and methods of the home environment, because the home environment has characteristics that are very different from those of the hospital in terms of medical facility, human resources, the medical knowledge of operator, and other factors. Thus, the approach for simply building the same monitoring architecture as the hospital will dramatically increase the time and human resources necessary for healthcare services delivery. It may also be unacceptable for patients due to its obtrusive and stressing nature. Within the scope of continuity of healthcare, the need to move beyond passive monitoring to efficient mechanisms for personalized healthcare is becoming more and more evident [79]. As was discussed in previous subchapter, one of the emerging requirements for a healthcare system is to provide various health monitoring and disease detection services in the context of each user’s characteristics. The key goal of the healthcare system at a hospital is to detect anomalies in patients’ physiological parameters as accurately as possible and to make diagnoses of potential diseases based on the detected anomalies, whereas the objective of home healthcare is to provide simple and easy methods for monitoring patients’ health status while they carry out the activities of daily life [80].
The service scenario for personalized chronic disease care is shown in Fig. 2.21. When a patient comes to the hospital, the patient receives several medical examinations in order to check the patient’s health status and to check for possible chronic diseases. If a chronic disease is detected, a physician verifies the detection results as to whether or not they are valid. The patient-specific disease care services are then uploaded to Integrated Healthcare System and the system interacts with the physician. When the patient goes back home and monitors their health status using home medical equipment, the measured data are transferred to an informant at a remote site. The informant monitors the patient’s health data and periodically stores the data in the medical database in the hospital. If an anomaly is detected in the patient’s health data, the informant reports the anomaly to the physician. Then, the physician examines the patient’s health data and makes a diagnosis. If necessary, the patient goes to the hospital for further examination.
2.4.2 Cloud-based Personalized Healthcare Service Scenario

The service scenario for personalized chronic disease care in cloud environment is shown in Fig. 2.22. When a patient comes to a hospital, the patient receives several medical examinations in order to check their disease status and to provide the appropriate treatment. The patient’s profile and health data are stored in a personalized virtual machine in the hospital cloud, and the personalized virtual machine is synchronized to the home healthcare system. When the patient goes back home and monitors their health status using the portable devices, the measured data are transferred to a mobile device and delivered to the home healthcare system. The patient-specific chronic disease care services are uploaded to the home healthcare system by the healthcare system administrator in the hospital and the system interacts with the physician and an informant at a remote site. The patient’s mobile device shares the sensing data with other mobile devices and with the home healthcare system in home environment. The informant monitors the patient’s health data and periodically stores the data in the medical database in the hospital. If an anomaly is detected in the patient’s health data, the informant reports the anomaly to the physician. The physician then examines the patient’s health data and makes a diagnosis. If necessary, the patient goes to the hospital for a closer examination. In the hospital cloud, the service broker system allocates resources in hospital cloud to the patient’s home cloud virtual machine. The broker system divides a workflow of chronic disease care service into tasks and selects cloud resources to execute the partitioned tasks. During the resource selection process, service levels of patients are considered. Also, the selected cloud resources are used to build virtual machines for home healthcare environment.
Figure 2.22: Service scenario for the personalized chronic disease management in home-hospital clouds.
2.5 Predicted Cost Model for Integrated Healthcare Service

We present the predicted cost model for the proposed scenarios of integrated healthcare service in home-hospital environment using Markov process. Predicting cost in the health care environment is a challenging dilemma for medical professionals. The importance of a viable cost model incorporating outcomes measurement and payment schemes is of interest. Healthcare administrators want to assure that the delivery of services is appropriate as identified by federal government guidelines, rules and regulations. A critical starting point is to provide the framework necessary to provide a cost model that considers the general factors of healthcare encounters, patient diagnosis, treatment and the related costs that can be used to describe this complex problem. The very stochastic nature of disease treatment can lead to substantial variation in experience between and among classes of enrollees, their diseases, and treatment utilization patterns. The most common approach to analyzing cost of disease is the traditional method of summing the number of events occurring in the system over a period of time and calculating the mean and a standard deviation of cost. There is a need for more sophisticated models to predict cost in the healthcare environment. A wide variety of conceptual and statistical models exist, both deterministic and stochastic, to measure utilization in health research. The deterministic models are traditional model (summing of events), and decision analysis (decision tree). The limitation of the traditional model is its inability to account for non-symmetric aspect of cost data and the lack of consideration of the utilization patterns of the population. Decision trees can be effective models in economic and policy analyses, because they can provide information to patients and practitioners about risk and cost. The difficulties with this model arise when timing becomes a concern. This problem becomes apparent when the time interval is several years or there are repeated events in a shorter time interval [93].

To resolve the limitations above, the Markov process has been widely used for modeling of epidemic progression of various diseases such as influenza, tuberculosis and HIV, and the pathways for subjects to predict utilization or possible pathways through the healthcare system [94-98]. Among them, to predict utilization in various healthcare plans and healthcare systems, Kapadia et al. studied 305 patients at a 90 bed comprehensive rehabilitation hospital in a major metropolitan area over a six months period [98]. The authors used hospital service charges and diagnosis to measure the utilization patterns of the patients. Also, Beland studied ambulatory care in Montreal, Canada. The author used the physician claims from clinic visits, hospitalization and emergency room visits and adopted a Markov chain to predict utilization to show the differences between
population demographics such as age and gender. The author showed that the corresponding changes in the traditional model of counting visits to physicians can be modeled using Markov process. The results of the above literature review indicate that the Markov process is appropriate to estimate the utilization of a population of patients or enrollees. The Markov process can illustrate the difference in the treatment utilization patterns due to predictor variables such as gender and age [97]. Also, Leviton et al. examined the application of a generalized Markov process seems appropriate to predict utilization for patients with chronic or acute diseases [99].

In accordance with the previous literature, in this subchapter, we develop a Markov process-based cost prediction model for integrated healthcare system in home-hospital environment. Table 2.11 shows the states of integrated healthcare system in home-hospital environment used in this subchapter to develop the Markov process. The states represent typical disease management scenario of integrated healthcare system [93].

<table>
<thead>
<tr>
<th>State 0 ($m_0$)</th>
<th>No identified chronic disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1 ($m_1$)</td>
<td>Self-management at home</td>
</tr>
<tr>
<td>State 2 ($m_2$)</td>
<td>Consult an informant about abrupt change on at-home medical examination results</td>
</tr>
<tr>
<td>State 3 ($m_3$)</td>
<td>Visit hospital (periodic visit or abrupt disease state change)</td>
</tr>
<tr>
<td>State 4 ($m_4$)</td>
<td>Diagnosis by physician</td>
</tr>
<tr>
<td>State 5 ($m_5$)</td>
<td>Laboratory tests</td>
</tr>
<tr>
<td>State 6 ($m_6$)</td>
<td>Discharged from the hospital</td>
</tr>
<tr>
<td>State 7 ($m_7$)</td>
<td>Complete cure</td>
</tr>
</tbody>
</table>

The Markov process is being employed as the first component of this model to predict utilization for this healthcare problem. The transition or change in utilization from state $i$ to state $j$ is influenced by the prior state is denoted as:

$$P_{ij} = P[X_{n+1} = j \mid X_n = i]$$

(2.9)

, where $P_{ij}$ is the probability of going from state $i$ to state $j$ in one step or one increment in the time unit.
Due to the patient’s outpatient flow is managed by predetermined patient management care plan of hospital, it is not necessary to consider all the previous transitions, when determining next transition. All the transitions from the resident states require preset exit criteria and only the previous state influences the opportunity for transition to an alternate state. This means that being in a resident utilization state two transitions earlier is irrelevant to the state you are in presently. All the information that is needed is the previous transition state. This is using the memory less or Markovian property. The resident states of transition based on utilization for the Markov process will be defined as follows: The states defined in Table 2.7 lead to an eight state Markov chain with an absorbing barrier (State 7) and the resulting one step transition probability matrix is

\[ P_{77} = (P_{ij}) \], where \( i = 0, 1, 2, \ldots, 7 \) and \( j = 0, 1, 2, \ldots, 7 \) \hspace{1cm} (2.10)

We specify the time interval unit for the probability transition matrix for utilization to be one unit time \( T \). Consider the finite number of possible transitions for individuals in the model. Denote by the element

\[ P_{00} = \Pr \{ X_n = 0 \mid X_{n-1} = 0 \}, \text{ for any } n. \] \hspace{1cm} (2.11)

, where the probability that an individual starting in \( m_0 \) stayed in \( m_0 \) after one time period.

Similarly, the probabilities for staying in the same state after one transition are denoted as follows:

\[ P_{00}, P_{11}, P_{22}, P_{33}, P_{44}, P_{55}, P_{66}, P_{77} \] \hspace{1cm} (2.12)

Let \( f_{ii}^n = \Pr \{ X_n = i, X_v \neq i, v = 1, 2, \ldots, n-1 \mid X_0 = i \} \] \hspace{1cm} (2.13)

be the probability that, starting from state \( i \), the first return to state \( i \) occurs at the \( n \)th transition [100]. There are 64 possible transitions for an individual in this model. State \( m_7 \) (complete cure) is an absorbing barrier or state, defined as a state that once entered cannot be exited [100]. This means that the probabilities of transition from \( m_7 \) to states \( m_0 \) through \( m_6 \) are zero and the probability of starting in state seven and staying in state seven is one. Hence in the probability transition matrix Eq. (2.10) becomes

\[ P_{70} = P_{71} = P_{72} = P_{73} = P_{74} = P_{75} = P_{76} = 0 \text{ } & \text{ } P_{77} = 1 \] \hspace{1cm} (2.14)

The matrix can be partitioned into 4 subsets. The set of transition probabilities, \( \{P_{ij}\} \), has the following two properties:

(1) it has a period of one and

(2) since \( f_{77}^n = 1 \), it is positive recurrent.

Combining the previous two properties leads to the conclusion that \( \{P_{ij}\} \) is an ergodic set. This set
will be represented by the submatrix, \( E_{1\times 1} \) and defined as follows:

\[
E_{1\times 1} = \{ P_{i\gamma} \}.
\]  

(2.15)

The next submatrix of the partitioned matrix to be considered is the vector of zeroes,

\[
O_{1\times 7} = \{ P_{i\gamma} \mid i = 0, 1, 2, 3, 4, 5, 6 \}.
\]  

(2.16)

Once in the absorbing state the individual cannot leave the state; hence all the cell entries are zero.

The third partitioned submatrix to be defined will be the transient states,

\[
M_{7\times 7} = \{ P_{i\gamma} \mid i = 0, 1, 2, 3, 4, 5, 6 \} \quad \& \quad j = 0, 1, 2, 3, 4, 5, 6 \}.
\]  

(2.17)

The submatrix \( M \) includes all the transient states of the Markov chain. The probability of the first return, Eq. (2.13), for these states \((m_0, m_1, m_2, m_3, m_4, m_5, m_6)\) is less than one. The final submatrix to consider is the transition from a transient state to the absorbing state. This will be defined as

\[
L_{7\times 1} = \{ P_{i\gamma} \mid i = 0, 1, 2, 3, 4, 5, 6 \}
\]  

(2.18)

An alternative form of the probability transition matrix can now be illustrated with dimensions of partitioned matrices:

\[
E_{1\times 1}, O_{1\times 7}, L_{7\times 1}, M_{7\times 7}. \quad \text{It should be noted that the matrix } \ E_{1\times 1} \quad \text{is equivalent to the identity matrix,} \ I_{1\times 1}. \quad \text{Replacing } \ E_{1\times 1} \quad \text{with } \ I_{1\times 1} \quad \text{in the matrix results in the following}
\]

\[
P = \begin{pmatrix}
L_{7\times 1} & O_{1\times 7} \\
L_{7\times 1} & M_{7\times 7}
\end{pmatrix}
\]  

(2.19)

Determining the mean time or number cycles an individual occupies in a resident state requires some knowledge of linear algebra and the development of the fundamental matrix for Markov chain with an absorbing state. Kemeny and Snell [101] developed methodology for finding the mean time in each resident state before transition into the absorbing state. They proposed the following.

Let \( B_{m} \) be a square matrix raised to the power \( m \). If \( B^n \to 0 \) as \( m \to \infty \), then \( (I - B) \) has an inverse, and

\[
(I - B)^{-1} = I + B + B^2 + \ldots = \sum_{i=0}^{\infty} B^i.
\]  

(2.20)

For any Markov chain with an ergodic set, let the matrix \( M \) correspond to the set of transient states, as in Eq. (2.19). Then \( (I - M) \) has an inverse, and

\[
(I - M)^{-1} = I + M + M^2 + \ldots = \sum_{i=0}^{\infty} M^i
\]  

(2.21)
Substituting the matrix \( M \) from Eq. (2.19) into Eq. (2.20) proves Eq. (2.21).

Let \( N = (1 - M)^{-1} \) be the fundamental matrix for a Markov chain with an ergodic state [101]. The next consideration is the number of times for an individual that a transient state is occupied. Define \( \eta_{ij} \) to be the function assigning the total number of times that the process is in state \( m_j \) after starting from state \( m_i \) (restricting the choices to transient states, \( \{m_j \mid j = 0, 1, 2, 3, 4, 5, 6 \} \)). This quantity will be will be expressed as the sum of indicator variables

\[
\mu_{ij}^k = \begin{cases} 
0, & \text{if the process is in state } m_j \text{ after } k \text{ steps} \\
1, & \text{otherwise}
\end{cases}
\] (2.22)

Determining the expectation of the number of cycles an individual stays in a resident transition state, conditional on having just entered the system, follows [101]. The mean number of days spent in \( m_j \) after starting in state \( m_i \) is \( N_{ij} = E[\eta_{ij}] \) as can be seen from the following argument. It should be observed that

\[
\eta_{ij} = \sum_{k=0}^{\infty} \mu_{ij}^k. 
\]

Hence, \( E[\eta_{ij}] = E\left[ \sum_{k=0}^{\infty} \mu_{ij}^k \right] = N_{ij} \). Note that the \( \mu_{ij}^k \) the \( l, j \) element of \( M^k \). Here \( \eta \) is the matrix whose \( l, j \) element is \( M^k \). Then

\[
E[\eta] = \sum_{k=0}^{\infty} E[\mu_{ij}^k] = \sum_{k=0}^{\infty} M^k = N
\] (2.23)

Denote the expected numbers of days in seven transient states by \( T' \), taken from the proper row of \( N \).

The notation for the cost function is Eq. (2.24). Define the fixed cost to be a column vector, where each element of this \( 8 \times 1 \) matrix is the averaged costs per utilization state of the system. It should be noted that the elements for the states "no use of services" and "Complete cure" have no allowable costs associated with them. The cost function can be represented as Eq. (2.24)

\[
C_{8 \times 1} = \{0, c_1, c_2, c_3, c_4, c_5, c_6, 0\} \quad (2.24)
\]

The model is defined by multiplying \( T' \) and (2.16) with the result

\[
F(x_i) = T'_{8 \times 1}(x_i)C_{8 \times 1} \quad (2.25)
\]

, where \( x_i \) is the conditions of interest (gender, age, and diagnosis). The value of the function \( F \) is the predicted cost for an individual. The vector of utilization, \( T' \), has a dimension of \( 1 \times 8 \). The vector of cost, \( C \), has a dimension of \( 8 \times 1 \). Taking their product generates a scalar value, \( F_{1 \times 1} \), which is the predicted cost given gender, age, and diagnosis.
2.6 Summary

In this chapter, we investigated the characteristics of chronic disease care cycles and the prevalence of major chronic diseases. Among the diseases, we further investigated the concept and diagnostic criteria of MS. We also presented the state-of-arts regarding the healthcare system for caring chronic disease effectively. Then, based on the survey results, we identified the major problems regarding the diagnostic criteria for MS and the conventional healthcare systems. The investigated problems regarding the definition of MS are different importance among MS risk factors, difficulty in identifying potential MS patients, and inability to effectively describe the temporal changes of the status of MS. Also, the problems and limitations of the conventional healthcare systems for chronic disease care are summarized as disconnected environments among homes and hospitals, requirement for large computational performance, and patients-centric customizability. Then, we presented Markov process-based cost model to predict costs based upon healthcare service scenario of home-hospital integrated environments.

The proposed cost model has following limitations. If the number of transitions is small in one or more resident states with the addition of one or more resident states becoming ergodic, then an unstable probability transition matrix is generated. The unstable matrix cannot provide appropriate estimates. The proposed model has not been tested on a large administrative database of claims. However, during the test phase of the proposed model, the following issues need to be considered. The first issue is that the restrictions due to the database with the lack of demographic identifiers. The absence of ethnicity and marital status may lead to questions about the changes in utilization patterns for these groups of enrollees. A second issue about evaluation is the choice of deleting the multiple events per day for an individual. This may cause the states remaining transient in the probability transition matrices. Therefore, the issues above need to be considered during further evaluation of the proposed model.

To resolve the investigated problems in this chapter, we will propose novel MS risk quantification model and temporal progress control model to help physicians and patients with proactive diagnosing and effective controlling the MS status in Chapter 3 and 4. Then, we will propose a cloud-based personalized healthcare system architecture to manage chronic disease effectively using the proposed MS risk quantification and control model in Chapter 5 along with prototype implementation results.
Chapter 3. Diagnostic Decision Method for Metabolic Syndrome Based on Risk Quantification Model

This chapter proposes a novel MS risk quantification model for identifying and diagnosing potential MS patients. The risk quantification model is based on ASD analysis using a new weighted radar chart method. We evaluate the validity of the proposed model using a large clinical database and show the clinical effectiveness of the risk quantification model. We further elaborate the proposed model to resolve various theoretical and clinical issues such as comparison with the conventional MS diagnosis method, clinical benefits and applicability to other populations, differentiation of individuals having the same risk by the proposed risk quantification model, and determination of the reference ASD threshold values for identifying potential MS patients.

3.1 Introduction

Metabolic syndrome refers to a clustering of specific cardiovascular disease risk factors whose underlying pathology is thought to be related to insulin resistance. The risk factors include insulin resistance, obesity, dyslipidemia, and hypertension and it is known to increase the risk for cardiovascular disease and type II diabetes. Since metabolic syndrome helps to identify individuals at high risk for both cardiovascular disease and type II diabetes, it has become a major public healthcare issue in many countries. There has been much effort to establish diagnostic criteria for metabolic syndrome, but the current diagnostic criteria of metabolic syndrome have weaknesses, such as binary decision based on diagnostic criteria, equal weight among risk factors, and difficulty in estimating the temporal progress of the risk factors. To resolve these problems, this chapter proposes a risk quantification model for metabolic syndrome, which is based on areal similarity degree analysis between weighted radar charts comprising metabolic syndrome diagnostic criteria and examination results of risk factors. The clinical effectiveness of the proposed model is extensively evaluated by using data of a large number of subjects obtained from the third Korea National Health and Nutrition Examination Survey. The evaluation results show that the proposed model can quantify the risk of metabolic syndrome and effectively identify a group of subjects who might be classified into a potential risk group for having metabolic syndrome in the future.
3.1.1 Problem Description on Current Diagnostic Decision Method for Metabolic Syndrome

There have been many studies to investigate the risk factor structure of MS, the prevalence of MS, the impact of the risk factors in the development of MS, and the relationship between MS and other diseases. Some of these studies are found in [3-10]. Also, there is another approach to investigate the mechanism of MS from the microscopic perspective at the cell and mitochondria level [11-13]. Those previous studies have mostly aimed to analyze the effects of individual risk factors on MS, but due to the clustering characteristics of MS, it is necessary to take a holistic approach for a more accurate understanding of MS. Multivariate analysis has shown that the individual risk factors comprising the syndrome each carry different odds ratios for predicting the prevalence of CHD, the incidence of CHD, and CVD mortality. In addition to hyperglycemia, low HDL-C and hypertension usually indicate a significantly greater risk compared with the presence of obesity or high triglycerides [14][15]. The studies also illustrate another likely shortcoming of the current approach to diagnosing MS. Both the ATP III and WHO definitions weigh each risk component equally, yet it is clear that some risk factors included in the definitions have greater CVD predictive value than others. This fact is highlighted in other algorithms used to predict CVD risk using regression coefficients to assign different weights to risk factors [16][17], and it is apparent from studies that examined the risk of CVD in persons with one or two components of MS versus three or more [18].

According to these studies, MS should be considered a progressive process that over time leads to major complications. Its expression depends heavily on age and exposure to an unhealthy lifestyle. Diagnosis of MS should not be managed as an acute infection, and a binary decision based on a yes or no approach is not valid in chronic degenerative disorders such as MS. MS incorporates clinical entities in which insulin resistance plays a major pathological role. This concept results in a group of subjects with a spectrum of long-term risk for having the final outcomes. The risk depends on the characteristics of the population being examined. Future adaptations to the current definitions of MS should take into account ethnic variability. Selection of diagnostic thresholds should be supported by the risk for developing major complications, especially if treatment can prevent these outcomes. The following case is a typical example to show the weakness of current MS diagnostic criteria. A 55-year old man who had a BMI of 27 kg/m2, a WC of 98 cm, HDL-C of 36
mg/dl, triglycerides of 180 mg/dl, fasting glucose of 90 mg/dl, post-challenge 2 hours glucose of 180 mg/dl, fasting insulin of 32 μU/ml, and a blood pressure of 120/80 mmHg is not determined to have MS by either the NCEP or the WHO definition, despite the presence of insulin resistance, impaired glucose tolerance, hypertriglyceridemia, low HDL-cholesterol, and overweight [2][4]. Therefore, it is necessary to develop a method to resolve the binary nature and different importance of risk factors of MS diagnostic criteria.

There have been several studies for developing personalized disease diagnosis methods to identify and quantify patients’ health status, particularly in relation to chronic diseases [19][20][25]. Among them, Jeong et al. [20], proposed a novel patient status classification method (PSCM) to quantify the chronic disease status of patients, which is based on patient tier classification and radar chart priority calculation using surface measure of overall performance (SMOP) theory [21]. The PSCM model for patients with chronic diseases offers automatic medical service procedures in the form of an effective medical information visualization system. The PSCM process contains three parts: the patient tier classifier, the disease & complications identifier, and health risk quantification. Although the radar chart approach (RCA) and the SMOP method used in PSCM are very effective means of identifying the characteristics of patients status having chronic disease, it is known that these methods have the following weaknesses [22]:

- Equally weighted indicators which are problematic and unjustified
- No explanation for status levels and changes of status
- No information about risk quantification by comparing measurement results with reference criteria

Therefore, to resolve the shortcomings of MS diagnosis methods discussed earlier, this chapter first proposes a new weighted radar chart construction method which extends the conventional radar chart-based chronic disease classification method to quantitatively describe the status of MS risk factors. Then, it presents a novel MS risk quantification model that is based on the similarity analysis between the proposed weighted radar charts consisting of MS diagnostic criteria and the examination results of MS risk factors of a subject. In the following subchapters, a description of the proposed model for quantifying MS risk of a subject is presented, along with its related weighted radar chart construction procedures. Furthermore, the effectiveness of the proposed model is extensively evaluated using data from the third Korea National Health and Nutrition Examination Survey (KNHANES III) [27] while emphasizing the capability to identify potential MS patients. Finally, the discussion and open issues are presented.
3.2 Related Work on Radar Chart Model

Radar charts consist of a set of performance indicators which are set in a circle. The indicators are typically standardized from zero to one, one indicating the highest possible performance. The performance degrees of all indicators are used to construct a plot for the whole system. The quality of the radar charts depends on the validity, reliability and comprehensiveness of the performance indicators. It is known that radar charts have two important advantages, namely, self-evident visualization and overall performance measurement [22]. Due to these advantages, radar charts are popularly used to assess and quantify the performance of various evaluation objectives and to present visual comparison of performance in various fields, especially business management. Several previous related studies in various disciplines have been carried out. Pati et al. presented rational expressions to quantify building performance using radar charts [31], whereas Kaczynsk et al. provided a framework using radar charts to qualitatively evaluate the impact of blended learning in the learning environment [32]. Chang et al. presented an adaptation of the radar chart to analyze the quality of service and quality of experience of several computer network services. The authors demonstrated the usability of the radar chart method in application performance analysis, application recommendation, and network diagnosis [34]. Schmid et al. used radar charts to benchmark national labor markets in the European Commission [22]. Shrestha developed a social scientific model to measure gendered share of resources between males and females using a radar chart and presented the applicability of the proposed methodology to show existing gender inequality in society in a variety of areas [35].

Radar charts have rarely appeared in biomedical engineering and healthcare literature. However, several precursors can be found. Saary demonstrated the usefulness of radar charts as an effective method to convey the meaning and analysis results of multivariate healthcare data to audiences through graphical visualization [36]. Pitt et al. investigated the use of information graphics to display the outputs of health technology assessment in the United Kingdom through radar chart representation and proposed a framework for developing effective information graphics in health technology assessment [37]. Agostinelli et al. evaluated the effects of the various graphical display formats e.g., bar, radar, and linear wave charts on understanding of hospital ward data [40]. Björkegren et al. performed a controlled study of general symptom reporting in fibromyalgia patients and evaluated the prevalence of 30 general symptoms experienced during the study period by using radar charts [38]. Chanques et al. compared the feasibility, validity and performance of five self-reported in-
tensity measurement scales for pain assessment in critically ill patients. The classification and evaluation of pain scales was done by using the radar chart [39]. In the disease diagnosis and management field, there have been several studies for developing personalized disease diagnosis methods to identify and quantify patients’ health status, particularly in relation to chronic diseases [19][20][25]. Among them, Jeong et al. [20], proposed a PSCM to quantify the chronic disease status of patients by using patient tier classification and radar chart priority calculation [21], which showed the applicability of the radar chart method in preliminary diagnosis of chronic disease. The proposed PSCM for patients with chronic diseases offered automatic medical service procedures in the form of an effective medical information visualization system.

There have been several previous studies to improve shortcomings of the traditional radar chart method. In [41], the authors proposed a radar chart-based fusing method for multidimensional environmental parameters to transform the linguistic description of environmental comfort levels into linear and nonlinear membership functions. Hongliang et al. presented an improved radar chart model based on transforming the conventional radar chart into a polar coordinates-based improved radar chart. The improved radar chart allows different weights on the indicators [42]. Also, the authors of [33] analyzed various vulnerabilities of systems and proposed classification patterns for the investigated vulnerabilities based on life cycles. The factors influencing the vulnerabilities were assessed using the radar chart with weights on the factors. According to the literature survey, the previous work has been mainly focused on improving the traditional radar chart to allow different weights on the indicators. However, to apply the radar chart method to risk quantification of MS diagnostic criteria, other limitations of the traditional radar chart should be resolved.

Therefore, to resolve the shortcomings of MS diagnosis methods discussed earlier though the radar chart methodology, this chapter proposes a new weighted radar chart construction method which extends the conventional radar chart-based chronic disease classification method to quantitatively describe the status of MS risk factors. Then, it presents a novel MS risk quantification model that is based on the similarity analysis between the proposed weighted radar charts consisting of MS diagnostic criteria and the examination results of MS risk factors of a subject. In the following subchapters, a description of the proposed model for quantifying MS risk of a subject is presented, along with its related weighted radar chart construction procedures. The effectiveness of the proposed model is evaluated using data from KNHANES III [27] while emphasizing the capability to identify potential MS patients. Finally, the discussion and open issues are presented.
3.3 Novel MS Risk Quantification Model

3.3.1 Construction of Weighted Radar Chart for Describing the Status of MS Risk Factors

As we discussed in the previous subchapter, the radar chart is a very useful method for qualitative data analysis despite its weaknesses, and it is useful for preliminary identification of a patient’s chronic disease and disease status [20]. Therefore, to mitigate the weaknesses of the radar chart, we have adopted the analytic hierarchy process (AHP) in determining the different weights of MS risk factors. Also, we propose a method for constructing a weighted radar chart. AHP is a well-known multi-criteria decision making method developed by Saaty [23], which allows decision makers to model a complex problem in a hierarchical structure, investigating the relative importance of components comprising the problem. To determine the weights of each risk factor of MS by using AHP, we first decompose the problem, in which the risk factors should be determined, into a hierarchy of more easily comprehended sub-problems and obtain the relative importance between each risk factor by pairwise comparison. By constructing pairwise comparison matrix and checking the consistency, then, weights can be computed. Thus, we apply AHP to determine the weights of each risk factor of MS.

Many variables are put into the same coordinate plane, the area is the representation function for the whole quality, and the shape gives the detail characteristics. In the weighted radar chart, every input variable value is expressed by radial $r_i$ of a unit circle, and $w_i$ is the weight coefficient. The unit circle could be divided into $n$ parts according to different weight coefficient $w_i$ and the sum of the $w_i$ coefficient is equal 1. On the circle, the $n$ rays represent the $n$ input variables, and the $r_i$ measures value of an input variable that falls in a relevant ray. Connecting the points $r_i$, which corresponds to measured values of different inputs, a weighted radar chart could be obtained. Since input data for each MS risk factor in the weighted radar chart have different measurement scales, the input data are normalized to fit into the new input range from 0 to 1.

Let the result of the $i_{th}$ risk factor be $x_i$ and the maximum and minimum value of input data be $x_{i_{\text{max}}}$ and $x_{i_{\text{min}}}$. Then, $x_{i_{\text{norm}}}$, the normalized value of $x_i$ can be written as

$$x_{i_{\text{norm}}} = \frac{x_i - \min_{i_{\text{min}}}}{x_{i_{\text{max}}} - \min_{i_{\text{min}}}}$$  \hspace{1cm} (3.1)

Since each risk factor has a different effect on MS status, it is necessary to separately determine the weight of each risk factor to the disease status. To separately decide the relative weight of the risk factors, we utilized the results of risk factor analysis for MS. Fig. 3.1 shows the decomposition of the risk factor structure.
for MS. In this subchapter, we use the following sample prevalence results of each risk factor. A description of how to determine the prevalence of risk factors will be provided in Subchapter 4.

\[
[F_{FG}, F_{WC}, F_{HDL-C}, F_{TG}, F_{BP}] = [0.281, 0.265, 0.452, 0.381, 0.456]
\]  

(3.2)

Figure 3.1: Decomposition of risk factor structure for MS.

Then, pairwise comparison matrix A for risk factors using the sample prevalence can be setup as follows [23]:

<table>
<thead>
<tr>
<th>FG</th>
<th>WC</th>
<th>HDL-C</th>
<th>TG</th>
<th>BP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0603</td>
<td>0.6217</td>
<td>0.7375</td>
<td>0.6162</td>
</tr>
<tr>
<td>0.9431</td>
<td>1</td>
<td>0.5863</td>
<td>0.6955</td>
<td>0.5811</td>
</tr>
<tr>
<td>1.6085</td>
<td>1.7057</td>
<td>1</td>
<td>1.1864</td>
<td>0.9912</td>
</tr>
<tr>
<td>1.3559</td>
<td>1.4377</td>
<td>0.8429</td>
<td>1</td>
<td>0.8355</td>
</tr>
<tr>
<td>1.6228</td>
<td>1.7208</td>
<td>1.0089</td>
<td>1.1969</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
A = \begin{bmatrix}
1 & 1.0603 & 0.6217 & 0.7375 & 0.6162 \\
0.9431 & 1 & 0.5863 & 0.6955 & 0.5811 \\
1.6085 & 1.7057 & 1 & 1.1864 & 0.9912 \\
1.3559 & 1.4377 & 0.8429 & 1 & 0.8355 \\
1.6228 & 1.7208 & 1.0089 & 1.1969 & 1
\end{bmatrix}
\]

where \(a_{ij}\) is the relative importance of the \(i_{th}\) element in the \(j_{th}\) indicator criterion level in terms of its contribution to the disease status, and \(n\) is the rank of this matrix. \(a_{ij}\) is calculated as a ratio of \(F_i\) to \(F_j\) listed in Eq. (3.2).

Once the pairwise comparison matrix has been established, the weight of each element being compared can be calculated. In this chapter, we have used the logarithmic least-square method to obtain the weights. The relative weight vector \(W\) can be obtained by solving the following equations [23]:

- 69 -
\[
\text{Minimize } \sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij} - \frac{w_j}{w_i})^2 \\
\sum_{i=1}^{n} w_i = 1, \quad w_i > 0 \quad \text{for } i = 1, \ldots, n.
\]

Obtained vector \( W \) for the relative weights of risk factors is given as

\[
W = [\text{FG, WC, HC, TR, BP}] = [0.1531, 0.1444, 0.2463, 0.2076, 0.2485]. \quad (3.3)
\]

To check the consistency of the risk factor values in pairwise comparison matrix, a consistency ratio \((C.R.)\) is used to determine the degree of consistency. If \( C.R. \leq 0.1 \), it means that the consistency level is satisfactory. The \( C.R. \) and consistency index \((C.I.)\) are defined as follows.

\[
C.R. = \frac{C.I.}{R.I.}
\]

\[
C.I. = \frac{\lambda_{\text{max}} - n}{n - 1}
\]

where \( \lambda_{\text{max}} \) is the maximum eigenvalue of the pairwise comparison matrix. The random index \((R.I.)\) is shown in Table 3.2 [23].

<table>
<thead>
<tr>
<th>Matrix order (n)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random index (R.I.)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.58</td>
<td>0.9</td>
<td>1.12</td>
<td>1.24</td>
<td>1.32</td>
<td>1.41</td>
</tr>
<tr>
<td>Matrix order (n)</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Random index (R.I.)</td>
<td>1.45</td>
<td>1.49</td>
<td>1.51</td>
<td>1.48</td>
<td>1.56</td>
<td>1.57</td>
<td>1.59</td>
<td></td>
</tr>
</tbody>
</table>

By substitute variable above by numerical value, the \( C.R. \) of the pairwise comparison matrix can be calculated as follows. Since \( C.R. \) is less than or equal 0.1, the consistency level is acceptable.

\[
\lambda_{\text{max}} = 5.00, \quad n = 5, \quad C.I. = \frac{\lambda_{\text{max}} - n}{n - 1} = \frac{5.00 - 5}{5 - 1} = 0 \\
C.R. = \frac{C.I.}{R.I.} = \frac{0}{1.12} = 0.
\]

The computed weights and allocation of angle for MS risk factors are shown in Table 3.3. From the table, we can observe that the examination results of waist/hip ratio, triglyceride, and HDL-cholesterol con-
tribute the main factor of metabolic syndrome disease, which comply with the results of [3].

Table 3.3: Computed weights for each indicator and allocation of angle in radar chart.

<table>
<thead>
<tr>
<th>Examination Test Type</th>
<th>Weight (%)</th>
<th>Allocation of Angle (Weight(%)*360°) (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting Glucose (FG)</td>
<td>15.31</td>
<td>55.1</td>
</tr>
<tr>
<td>Waist Circumference (WC)</td>
<td>14.44</td>
<td>52.0</td>
</tr>
<tr>
<td>HDL-Cholesterol (HDL-C)</td>
<td>24.63</td>
<td>88.7</td>
</tr>
<tr>
<td>Triglycerides (TG)</td>
<td>20.76</td>
<td>74.7</td>
</tr>
<tr>
<td>Blood Pressure (BP)</td>
<td>24.85</td>
<td>89.5</td>
</tr>
</tbody>
</table>

The circle of the weighted radar chart is marked off in accordance with the number of risk factors and the weights calculated above. Some radial lines are formed by the center of the circle and the marked point. These lines are regarded as coordinate axes. We pretreat the data on these coordinate axes and connect the marked points. Then, polygons for the values of risk factors can be obtained. This is the weighted radar chart of MS risk factors status for a subject. Fig. 3.2 depicts a weighted radar chart using a sample subject’s examination results shown in Table 3.4 [24].

Table 3.4: Sample examination results of type 2 diabetes patient (male).

<table>
<thead>
<tr>
<th>Examination Test Type</th>
<th>Sample Patient</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting Glucose (FG) (mg/dl)</td>
<td>176</td>
<td>295</td>
<td>51</td>
</tr>
<tr>
<td>Waist Circumference (WC) (cm)</td>
<td>88</td>
<td>127</td>
<td>43</td>
</tr>
<tr>
<td>HDL-Cholesterol (HDL-C) (mg/dl)</td>
<td>50</td>
<td>103</td>
<td>15</td>
</tr>
<tr>
<td>Triglycerides (TG) (mg/dl)</td>
<td>164</td>
<td>929</td>
<td>31</td>
</tr>
<tr>
<td>Blood Pressure (SBP/DBP) (mmHg)</td>
<td>125/80</td>
<td>203/140</td>
<td>70/35</td>
</tr>
</tbody>
</table>

The normalization of examination results is performed by using Eq. (3.1), and the maximum and minimum values of each risk factor are imported from the investigated results in [27]. Among the risk factors, the status of HDL-cholesterol becomes worse as the value decreases. Thus, the result of HDL-cholesterol is normalized as
\[ x_{HDL-C_{new}} = 1 - \frac{x_{HDL-C} - x_{HDL-C_{min}}}{x_{HDL-C_{max}} - x_{HDL-C_{min}}}. \] (3.4)

Figure 3.2: Weighted radar chart for status of MS risk factors.

### 3.3.2 MS Risk Modeling Using Areal Similarity Degree

As we discussed in the Subchapter 3.2, the current diagnostic criteria of MS have the following weaknesses:

- Variation in the importance of risk factors on the development of CVD and T2DM are not considered
- Diagnosis is based on the number of risk factors exceeding the thresholds of each, so it is difficult to estimate the risk of MS for a person who is diagnosed as not having MS.
- It is difficult to manage the progressive process over time.

Therefore, in this subchapter, we propose an MS risk quantification model utilizing the comparison
result of two weighted radar charts constructed using MS diagnostic criteria and the examination results of a subject, respectively. The comparison is performed based on the areal similarity degree analysis defined in this subchapter.

Let us consider a partial radar chart consisting of two polygons $A_{ij}$ and $B_{ij}$ depicted in Fig. 3.3.

![A partial radar chart consisting of two polygons.](image)

**Figure 3.3**: A partial radar chart consisting of two polygons.

**Definition 3.1.** Let radar chart $R$ be a set of disjoint polygons $A_{ij}$, where $i=1,\ldots,n$ and $j=(i+1)\ mod\ n$.

**Definition 3.2.** Let polygon $A_{ij}$ in radar chart $R$ be a polygon consisting of vertices $O, A_i, A_j$, i.e., triangle $\Delta OA_i A_j$, where $\overline{OA_i} = r_i, \overline{OA_j} = r_j, \angle O = \theta_i, i=1,\ldots,n, j=(i+1)\ mod\ n$. And $r_i$ is the value of $i$th indicator.

**Definition 3.3.** Let areal similarity degree (ASD) of two polygons $A_{ij}$ and $B_{ij}$ be the ratio of intersection area of the two polygons $A_{ij}$ and $B_{ij}$ over the area of polygon $B_{ij}$. In other words, areal similarity degree of two polygons

$$S(A_{ij} \mid B_{ij}) = \frac{\text{Area of intersection of polygon } A_{ij} \text{ and } B_{ij}}{\text{Area of polygon } B_{ij}}.$$  

(3.5)

In Eq. (3.5), polygon $B_{ij}$ is used for the reference polygon determined by the thresholds of MS risk factors and polygon $A_{ij}$ is determined by the examination results of the risk factors for an individual. Thus, Eq. (3.5) describes how the individual’s examination results of the risk factors are close to the thresholds of MS risk factors.
**Definition 3.4.** Given two polygons $A_{ij}$ and $B_{ij}$, let polygon $A_{ij}$ include polygon $B_{ij}$, iff $r_{ia} \geq r_{ib}$ and $r_{ja} \geq r_{jb}$.

**Theorem 3.1.** Given two indicators $i$ and $j$, the areal similarity degree of two polygons $A_{ij}$ and $B_{ij}$ namely, $S(A_{ij}|B_{ij})$ can be calculated as follows.

$$S(A_{ij}|B_{ij}) = \begin{cases} 1, & \text{if } A_{ij} \text{ includes } B_{ij} \\ \frac{\text{Area of } A_{ij}}{\text{Area of } B_{ij}}, & \text{if } B_{ij} \text{ includes } A_{ij} \\ \frac{r_{ja} \cdot r_{ib} - Q}{r_{ja} \cdot r_{ib}}, & \text{where } Q = \frac{r_{ia} \cdot r_{ib} (r_{ja} - r_{jb})}{r_{ja} - r_{jb} + r_{ja} (r_{jb} - r_{ja})}, & \text{if } r_{ia} > r_{ib} \text{ and } r_{ja} > r_{jb} \\ \frac{r_{ja} \cdot r_{ib} - Q'}{r_{ja} \cdot r_{ib}}, & \text{where } Q' = \frac{r_{ja} \cdot r_{ib} (r_{ia} - r_{ib})}{r_{ia} - r_{ib} + r_{ja} (r_{ib} - r_{ia})}, & \text{if } r_{ia} < r_{ib} \text{ and } r_{ja} < r_{jb} \end{cases}$$  

(3.6)

**Proof.**

1) According to Definition 3.4, when $A_{ij}$ includes $B_{ij}$, the intersection of the two polygons is $B_{ij}$. Therefore, $S(A_{ij}|B_{ij})$ is 1.

2) Similarly, when $B_{ij}$ includes $A_{ij}$, the intersection of the two polygons is $A_{ij}$. Therefore, $S(A_{ij}|B_{ij})$ is \[
\frac{\text{Area of } A_{ij}}{\text{Area of } B_{ij}}. \tag{3.7}\]

3) In Fig. 4.4, let $\angle B_j = \alpha$. Then, by applying the law of sines [26], we can obtain

$$\frac{A_i A_j}{\sin \theta_i} = \frac{r_{i_a}}{\sin \alpha} \quad \text{and} \quad \sin \alpha = \frac{r_{i_a} \cdot \sin \theta_i}{A_i A_j}.$$  

Then, by applying Menelaus’ theorem [26], we can get

$$\frac{O_{B_j}}{B_j A_i} \cdot \frac{A_i P}{P A_j} \cdot \frac{A_j B_j}{B_j O} = -1.$$  

By substituting variables with values, we can get the following equations.

$$\frac{r_{i_a}}{r_{i_a} - r_{ib}} \cdot \frac{P_{B_j}}{P_{B_j}} \cdot \frac{r_{ia} - r_{ja}}{r_{ia}} = -1.$$
So, \( PB_j = \frac{r_{j_b} (r_{i_a} - r_{i_g}) \cdot PB_j}{r_{j_a} (r_{j_b} - r_{j_a})} \).

Since \( B_i B_j = PB_i + PB_j \), \( PB_j \) is calculated as follows.

\[
PB_j = B_i B_j \cdot \frac{r_{j_a} (r_{j_b} - r_{j_a})}{r_{j_a} (r_{i_a} - r_{i_g}) + r_{j_a} (r_{j_b} - r_{j_a})}.
\]

In Fig. 4.4, the area of \( \Delta A_j B_j P \) becomes

\[
\Delta A_j B_j P = \frac{1}{2} \cdot OA_j \cdot PB_j \cdot \sin \alpha
\]

\[
= \frac{r_{i_a} \cdot r_{i_g} (r_{j_b} - r_{j_a})^2}{2(r_{j_a} (r_{i_a} - r_{i_g}) + r_{j_a} (r_{j_b} - r_{j_a}))}.
\]

Let the area of \( \triangle OB_j B_i \) be \( U \), then \( U \) becomes

\[
U = \frac{1}{2} \cdot r_{j_b} \cdot r_{i_g} \cdot \sin \theta_i.
\]

Let the area of \( \Box OB_j PA_i \) be \( T \), then \( T \) becomes

\[
T = U - \Delta A_j B_j P
\]

\[
= \frac{1}{2} \cdot r_{j_b} \cdot r_{i_g} \cdot \sin \theta_i - \frac{r_{i_a} \cdot r_{i_g} (r_{j_b} - r_{j_a})^2}{2(r_{j_a} (r_{i_a} - r_{i_g}) + r_{j_a} (r_{j_b} - r_{j_a}))}
\]

\[
= \frac{1}{2} (r_{j_b} \cdot r_{i_g} - Q) \sin \theta_i, \quad Q = \frac{r_{i_a} \cdot r_{i_g} (r_{j_b} - r_{j_a})^2}{r_{j_a} (r_{i_a} - r_{i_g}) + r_{j_a} (r_{j_b} - r_{j_a})}.
\]

According to Definition 3.3, \( S(A_j | B_{ij}) \) becomes \( \frac{T}{U} \) and we obtain

\[
T = \frac{1}{2} (r_{j_b} \cdot r_{i_g} - Q) \cdot \sin \theta_i,
\]

\[
U = \frac{1}{2} r_{j_b} \cdot r_{i_g} \cdot \sin \theta_i.
\]

Therefore, we can get \( S(A_j | B_{ij}) \) as follows.

\[
S(A_j | B_{ij}) = \frac{r_{j_b} \cdot r_{i_g} - Q}{r_{j_b} \cdot r_{i_g}}, \text{where } Q = \frac{r_{i_a} \cdot r_{i_g} (r_{j_b} - r_{j_a})^2}{r_{j_a} (r_{i_a} - r_{i_g}) + r_{j_a} (r_{j_b} - r_{j_a})}.
\]  \hspace{1cm} (3.8)

4) Let \( \angle B_i = \alpha \) and in a similar way to case 3), we can obtain \( S(A_j | B_{ij}) \) for case 4) as follows:

\[
S(A_j | B_{ij}) = \frac{r_{j_b} \cdot r_{i_g} - Q}{r_{j_b} \cdot r_{i_g}}, \text{where } Q = \frac{r_{i_a} \cdot r_{i_g} (r_{j_b} - r_{j_a})^2}{r_{j_a} (r_{i_a} - r_{i_g}) + r_{j_a} (r_{j_b} - r_{j_a})}.
\]  \hspace{1cm} (3.9)

\hspace{1cm} (q.e.d.)

Since current MS diagnosis methods judge the presence of MS only by checking whether the examination results of each risk factor exceed thresholds of the risk factors, it is not possible to provide any infor-
mation when the examination results do not exceed the thresholds. Thus, when the examination results are within the thresholds of risk factors, the risk for a subject may be decided by the physician’s knowledge. ASD defined in Theorem 1 shows the overall achievement of examination results of a subject versus thresholds of MS risk factors. Let us assume that, in Fig. 3, the \( i_{th} \) indicator and the \( j_{th} \) indicator are two risk factors of MS, and data set A (solid line) and data set B (dashed line) are the examination results and thresholds of the two risk factors, respectively. When data set A incudes data set B, we can decide that both examination results exceed the thresholds. When either one or none of the examination results exceed a threshold, the intersection between polygons consisting of thresholds and examination results is calculated and is used to express the achievement for each risk factor, respectively. Thus, we can quantify the risk for those two risk factors by using the intersection.

**Definition 3.5.** Let the areal similarity degree of two radar charts \( R_1 \) and \( R_2 \) be the ratio of the intersection area of the two radar charts, \( R_1 \) and \( R_2 \), over the area of radar chart \( R_2 \). In other words, the areal similarity degree of two radar charts

\[
S(R_1 \mid R_2) = \frac{\text{Area of intersection of radar charts } R_1 \text{ and } R_2}{\text{Area of radar chart } R_2}.
\]  

(3.10)

**Theorem 3.2.** Let the areal similarity degree of two radar charts \( R_A \) and \( R_B \) be \( S(R_A \mid R_B) \). Then, \( S(R_A \mid R_B) \) is the weighted sum of \( S(A_{ij} \mid B_{ij}) \), where \( w_i \) is a weight of \( A_{ij} \) and \( B_{ij} \), and \( i=1, \ldots, n \) and \( j=(i+1) \mod n \), and \( w_i = \frac{\theta_i}{360} \).

In other words, \( S(R_A \mid R_B) = \sum_{i=1}^{n} \frac{\theta_i}{360} \cdot S(A_{ij} \mid B_{ij}) \), where \( \sum_{i} \theta_i = 360 \), \( j=(i+1) \mod n \).  

(3.11)

**Proof.**

Since a radar chart comprises \( n \) disjoint parts, the areal similarity of two radar charts can be calculated by adding the weighted areal similarity degrees of \( n \) disjoint parts.

According to the definition, we can obtain the following:

\[
S(A_{12} \mid B_{12}) = \frac{\text{Area of intersection of polygon } A_{12} \text{ and } B_{12}}{\text{Area of polygon } B_{12}} = \frac{A_{12} \cap B_{12}}{B_{12}}
\]
\[
S(A_{nl} \mid B_{nl}) = \frac{\text{Area of intersection of polygon } A_{nl} \text{ and } B_{nl}}{\text{Area of polygon } B_{nl}} = \frac{A_{nl} \cap B_{nl}}{B_{nl}}.
\]

Also, weights of \(A_{12}\) and \(B_{12}\) are calculated as follows.

weight of \(A_{12}\) = weight of \(B_{12}\) = \(w_1 = \frac{\theta_1}{360} = \sum_{i,j} B_{ij} = \sum_{i,j} A_{ij} \)

weight of \(A_{nl}\) = weight of \(B_{nl}\) = \(w_n = \frac{\theta_n}{360} = \sum_{i,j} B_{ij} = \sum_{i,j} A_{ij} \).

Therefore, according to the definition, we can obtain

\[
S(R_A \mid R_B) = \frac{\text{Area of intersection of radar charts } R_A \text{ and } R_B}{\text{Area of radar chart } R_B}
\]

\[
= \frac{A_{12} \cap B_{12} + \cdots + A_{nl} \cap B_{nl}}{\sum_{i,j} B_{ij}}
\]

\[
= \frac{S(A_{12} \mid B_{12}) \cdot B_{12} + \ldots + S(A_{nl} \mid B_{nl}) \cdot B_{nl}}{\sum_{i,j} B_{ij}}
\]

\[
= S(A_{12} \mid B_{12}) \cdot w_1 + \ldots + S(A_{nl} \mid B_{nl}) \cdot w_n
\]

\[
= \sum_{i,j} S(A_{ij} \mid B_{ij}) \cdot w_i.
\]

(3.12)

\(q.e.d.\)

As shown in Theorem 3.2, since a weighted radar chart can be divided into polygons constructed by MS risk factors, we can calculate the overall ASD by combining the ASDs of each polygon. Thus, we can quantify the risk for all of the risk factors by using the sum of intersections.

In this chapter, we use the diagnostic criteria defined by NCEP ATP III in 2001 [28]. Among the criteria, we have adopted the waist circumference cut-off value for Korean subjects that was proposed by the Korean Medical Association [29] and the fasting glucose cut-off value that was proposed by the American Diabetes Association [30]. The thresholds of five risk factors for Korean subjects are listed in Table 3.5.
Table 3.5: Thresholds of five risk factors of MS for Korean subjects.

<table>
<thead>
<tr>
<th>Examination Test Type</th>
<th>Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting Glucose (FG) (mg/dl)</td>
<td>≥ 100mg/dl</td>
</tr>
<tr>
<td>Waist Circumference (WC) (cm)</td>
<td>≥ 90 (for male)</td>
</tr>
<tr>
<td></td>
<td>≥ 85 (for female)</td>
</tr>
<tr>
<td>HDL-Cholesterol (HDL-C) (mg/dl)</td>
<td>&lt; 40 (for male)</td>
</tr>
<tr>
<td></td>
<td>&lt; 50 (for female)</td>
</tr>
<tr>
<td>Triglycerides (TG) (mg/dl)</td>
<td>≥ 150mg/dl</td>
</tr>
<tr>
<td>Blood Pressure (BP) (mmHg)</td>
<td>≥ 130/85mmHg</td>
</tr>
</tbody>
</table>

A subject is determined to have MS, if three or more of the subject’s medical examination results exceed the thresholds of risk factors in Table 3.5. Thus, by comparing the weighted radar chart of the subject with the weighted radar chart of thresholds of risk factors, it is possible to effectively determine whether the subject has MS. Furthermore, by calculating ASD between two weighted radar charts, we can determine whether the subject can be classified into a potential risk group for having MS in the future.

Fig. 3.4 shows the two weighted radar charts for the subject and MS thresholds. Examination results in Table 3.3 have been used to construct the weighted radar charts. According to the examination results, the subject is determined not to have MS because only two examination results (fasting glucose and triglycerides) exceed the thresholds. However, the subject has T2DM, and the examination result of triglycerides also indicates the need for caution. Using conventional MS diagnostic criteria, it is difficult to estimate the risk of this kind of subject. We can compute the ASD of this subject as 0.93 based on Theorem 3.2. The clinical value of the proposed risk quantification model based on ASD will be evaluated and discussed in the Subchapter 3.4.
3.3.3 Discussion of Weighted Radar Chart Model

Since the proposed method is based on the calculation of the ASD of two weighted radar charts that comprise five polygons each, the shapes of the polygons can change as the order that the MS risk factors are arranged within the weighted radar charts changes. Further, these changes may cause changes in the calculated ASD results. This subchapter investigates the effects of changing the order of risk factors to the results. Let us consider the following property of the weighted radar chart.

**Property 3.1.** The ASD of two polygons $A_{ij}$ and $B_{ij}$ has a symmetric property. Let $A_{ij}$ and $B_{ij}$ be triangles comprising the ordered vertices $O, A_i,$ and $A_j$ and $O, B_i,$ and $B_j,$ respectively. Let $A_{ji}$ and $B_{ji}$ be triangles comprising the ordered vertices $O, A_j,$ and $A_i$ and $O, B_j,$ and $B_i,$ respectively. Then, $S(A_{ij}|B_{ij}) = S(A_{ji}|B_{ji}).$
Proof.

According to Definition 3 and Fig. 3, it is obvious that the following equation holds:

\[ S(A_{ij} \mid B_{ij}) = \frac{\text{Area of intersection of } A_{ij} \text{ and } B_{ij}}{\text{Area of } B_{ij}} = \frac{\text{Area of intersection of } A_{ji} \text{ and } B_{ji}}{\text{Area of } B_{ji}}. \]  

(3.13)

(q.e.d)

Using the property, we can develop the following corollary.

**Corollary 3.1.** Let us consider two weighted radar charts with 5 indicators as shown in Fig. 3.5. According to Theorem 3.2, the ASD of radar charts A and B is represented as follows:

\[ S(R_A \mid R_B) = w_1 \cdot S(A_{12} \mid B_{12}) + w_2 \cdot S(A_{23} \mid B_{23}) + w_3 \cdot S(A_{34} \mid B_{34}) + w_4 \cdot S(A_{45} \mid B_{45}) + w_5 \cdot S(A_{51} \mid B_{51}). \]  

(3.14)

Let us assume that we change the order of indicators by interchanging the position of indicators. For simple analysis, we consider exchanging the position of indicator 1 with another indicator. Then, there are two cases.

1) Exchange with an adjacent indicator

Let us assume that we exchange indicator 1’s position with indicator 2’s position as shown in Fig. 3.6.

Then, the ASD of the modified weighted radar chart is represented as

\[ S(R'_A \mid R'_B) = w_2 \cdot S(A_{21} \mid B_{21}) + w_1 \cdot S(A_{13} \mid B_{13}) + w_3 \cdot S(A_{34} \mid B_{34}) + w_4 \cdot S(A_{45} \mid B_{45}) + w_5 \cdot S(A_{52} \mid B_{52}). \]

By applying Property 3.1, the ASD difference of \( S(R'_A \mid R'_B) \) and \( S(R_A \mid R_B) \) becomes

\[ \text{Diff of } S(R'_A \mid R'_B) \text{ and } S(R_A \mid R_B) \]

\[ = S(R'_A \mid R'_B) - S(R_A \mid R_B) \]

\[ = (w_2 - w_1) \cdot S(A_{12} \mid B_{12}) + w_2 \cdot S(A_{23} \mid B_{23}) + w_3 \cdot [S(A_{32} \mid B_{32}) - S(A_{51} \mid B_{51})] + w_1 \cdot S(A_{13} \mid B_{13}) - w_2 \cdot S(A_{23} \mid B_{23}). \]

(3.15)

2) Exchange with a non-adjacent indicator

Let us assume that we exchange indicator 1’s position with indicator 3’s position as shown in Fig. 3.7.

Then, the ASD of the modified weighted radar chart is expressed as

\[ S(R'_A \mid R'_B) = w_3 \cdot S(A_{32} \mid B_{32}) + w_2 \cdot S(A_{21} \mid B_{21}) + w_1 \cdot S(A_{14} \mid B_{14}) + w_4 \cdot S(A_{45} \mid B_{45}) + w_5 \cdot S(A_{53} \mid B_{53}). \]
The ASD difference of \( S(R_A^* \mid R_B^*) \) and \( S(R_A \mid R_B) \) becomes

\[
\text{Diff of } S(R_A^* \mid R_B^*) \text{ and } S(R_A \mid R_B) = S(R_A^* \mid R_B^*) - S(R_A \mid R_B) = (w_3 - w_2) \cdot S(B_{23}) + (w_2 - w_1) \cdot S(B_{12}) + w_5 \cdot [S(B_{35}) - S(A_{31})] + w_1 \cdot S(A_{14}) - w_3 \cdot S(A_{34}) \cdot B_{34}.
\]

(3.16)

Figure 3.5: Sample weighted radar charts.

Figure 3.6: Exchange the positions of indicator 1 and 2.
Figure 3.7: Exchange the positions of indicator 1 and 3.

Therefore, by using Corollary 3.1, it is possible to calculate the changed ASD value, when the order of the risk factors is changed. For instance, the ASD of the weighted radar charts in Fig. 3.8 (a) is calculated as 0.93. If the position of the FG risk factor is moved between HDL-C and TG as shown in Fig. 3.8 (b), the changed ASD value is calculated as 0.92. The effects of changing the order of risk factors to the ASD analyses will be further investigated in the discussion part of the Subchapter 3.4.

Figure 3.8: Calculation of ASD for two weighted radar charts.
3.4 Evaluation and Discussion

3.4.1 Characteristics of Study Subjects

This chapter is based on data obtained from the third Korea National Health and Nutrition Examination Survey (KNHANES III) among non-institutionalized civilians in the Republic of Korea, which was conducted by the Korean Ministry of Health and Welfare in 2005. This survey was a nationwide representative study using a stratified, multistage probability sampling design for the selection of household units. The survey consisted of the following 4 components: the Health Interview Survey, the Health Behavior Survey, the Health Examination Survey, and the Nutrition Survey [27]. A total of 34,145 individuals from these sampling frames were included in the health interview survey; among them, 25,161 subjects aged over 20 years were identified as potential participants in our study. We excluded those with incomplete data for the standardized analysis. This resulted in a final analytical sample of 5,355 subjects (2276 male, 3079 female), aged over 20 years.

The characteristics of the study population, as stratified by gender, are presented in Table 3.6. The proportion of female subjects was higher than male (57.47% vs. 42.53%). The mean ages of the male and female subjects were 47.22±14.61 and 46.99±15.62 years, respectively. The average BMI was 23.99±3.10 and 23.52±3.38 kg/m² for male and female subjects, respectively. The percentage of subjects with diabetes mellitus in the male group was higher than that in the female group (6.90% vs. 4.48%). Also, 22.98% of the male subjects had hypertension, whereas 14.68% of the female subjects did. In this chapter, we have classified total subjects into two subject groups by gender and further classified each subject group into three sub-groups by age: young (from 20 to 39 years old), middle-aged (from 40 to 64 years old), and old (more than 65 years old), respectively. Therefore, we use a total of six sub-groups for the evaluation of our proposed risk quantification model.
Table 3.6: Characteristics of study subjects.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>2276 (42.53)</td>
<td>3079 (57.47)</td>
</tr>
<tr>
<td>Age, years</td>
<td>47.22±14.61</td>
<td>46.99±15.62</td>
</tr>
<tr>
<td>20-39 (young)</td>
<td>760 (33.39)</td>
<td>1118 (36.31)</td>
</tr>
<tr>
<td>40-64 (middle aged)</td>
<td>1170 (51.41)</td>
<td>1449 (47.06)</td>
</tr>
<tr>
<td>65+ (old)</td>
<td>346 (15.20)</td>
<td>512 (16.63)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.99±3.10 Mn 14.76</td>
<td>23.52±3.38 Mn 15.03</td>
</tr>
<tr>
<td></td>
<td>Mn 23.52±3.38</td>
<td>Mn 23.52±3.38</td>
</tr>
<tr>
<td></td>
<td>Mx 43.39 Mn 14.76</td>
<td>Mx 43.39 Mn 15.03</td>
</tr>
<tr>
<td></td>
<td>Mn 23.52±3.38</td>
<td>Mn 23.52±3.38</td>
</tr>
<tr>
<td>Fasting glucose (mg/dl)</td>
<td>97.61±22.97 Mn 51.0</td>
<td>93.06±19.78 Mn 60.0</td>
</tr>
<tr>
<td></td>
<td>Mn 93.06±19.78</td>
<td>Mn 93.06±19.78</td>
</tr>
<tr>
<td></td>
<td>Mx 295.0 Mn 51.0</td>
<td>Mx 295.0 Mn 60.0</td>
</tr>
<tr>
<td></td>
<td>Mx 295.0 Mn 51.0</td>
<td>Mx 295.0 Mn 60.0</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>84.30±8.74 Mn 39.0</td>
<td>78.46±9.54 Mn 41.5</td>
</tr>
<tr>
<td></td>
<td>Mn 78.46±9.54</td>
<td>Mn 78.46±9.54</td>
</tr>
<tr>
<td></td>
<td>Mx 133.5 Mn 39.0</td>
<td>Mx 133.5 Mn 41.5</td>
</tr>
<tr>
<td></td>
<td>Mx 133.5 Mn 39.0</td>
<td>Mx 133.5 Mn 41.5</td>
</tr>
<tr>
<td>HDL-cholesterol (mg/dl)</td>
<td>42.44±10.15 Mn 15.0</td>
<td>47.10±10.88 Mn 18.0</td>
</tr>
<tr>
<td></td>
<td>Mn 47.10±10.88</td>
<td>Mn 47.10±10.88</td>
</tr>
<tr>
<td></td>
<td>Mx 103.0 Mn 15.0</td>
<td>Mx 103.0 Mn 18.0</td>
</tr>
<tr>
<td></td>
<td>Mx 103.0 Mn 15.0</td>
<td>Mx 103.0 Mn 18.0</td>
</tr>
<tr>
<td>Triglycerides (mg/dl)</td>
<td>156.08±110.81 Mn 31.0</td>
<td>114.71±75.23 Mn 29.0</td>
</tr>
<tr>
<td></td>
<td>Mn 114.71±75.23</td>
<td>Mn 114.71±75.23</td>
</tr>
<tr>
<td></td>
<td>Mx 929.0 Mn 31.0</td>
<td>Mx 929.0 Mn 29.0</td>
</tr>
<tr>
<td></td>
<td>Mx 929.0 Mn 31.0</td>
<td>Mx 929.0 Mn 29.0</td>
</tr>
<tr>
<td>Systolic BP (mm Hg)</td>
<td>122.82±16.06 Mn 70.0</td>
<td>116.44±18.50 Mn 73.0</td>
</tr>
<tr>
<td></td>
<td>Mn 116.44±18.50</td>
<td>Mn 116.44±18.50</td>
</tr>
<tr>
<td></td>
<td>Mx 203.0 Mn 70.0</td>
<td>Mx 203.0 Mn 73.0</td>
</tr>
<tr>
<td></td>
<td>Mx 203.0 Mn 70.0</td>
<td>Mx 203.0 Mn 73.0</td>
</tr>
<tr>
<td>Diastolic BP (mm Hg)</td>
<td>80.74±10.37 Mn 35.0</td>
<td>74.72±10.26 Mn 38.0</td>
</tr>
<tr>
<td></td>
<td>Mn 74.72±10.26</td>
<td>Mn 74.72±10.26</td>
</tr>
<tr>
<td></td>
<td>Mx 140.0 Mn 35.0</td>
<td>Mx 140.0 Mn 38.0</td>
</tr>
<tr>
<td></td>
<td>Mx 140.0 Mn 35.0</td>
<td>Mx 140.0 Mn 38.0</td>
</tr>
<tr>
<td>Diabetes mellitus¹, n (%)</td>
<td>157 (6.90)</td>
<td>138 (4.48)</td>
</tr>
<tr>
<td>Hypertension², n (%)</td>
<td>523 (22.98)</td>
<td>452 (14.68)</td>
</tr>
</tbody>
</table>

BMI, body mass index; BP, blood pressure

Values are means ± SD or n (%); Mn, minimum; Mx, Maximum

¹Fasting glucose ≥126 mg/dl and/or physician-diagnosed diabetes mellitus

²Systolic BP ≥140 mm Hg and/or diastolic BP ≥90 mm Hg and/or physician-diagnosed hypertension
First of all, to analyze the relative weights of MS risk factors among each sub-group, we use prevalence-based weight determination for each MS risk factor for each sub-group [27]. For each sub-group, we counted the occurrence of examination results that exceeded the thresholds of each risk factor. Then, the number of occurrence of each risk factor was mapped into the prevalence result of each risk factor in Eq. (3.2). Table 3.7 shows the computed relative weights among MS risk factors. According to the analysis, the prevalence of lipid-related risk factors showed high occurrence, whereas fasting glucose and waist circumference showed relatively low occurrence in male subjects. For female subjects, HDL-cholesterol and waist circumference were identified as major risk factors, whereas triglycerides, which had the highest prevalence for male subjects, was not a major factor.

Table 3.7: Calculation of weights for MS risk factors among study subjects.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th></th>
<th>Female</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FG</td>
<td>WC</td>
<td>HDL-C</td>
<td>TG</td>
<td>BP</td>
<td>FG</td>
<td>WC</td>
</tr>
<tr>
<td>Young</td>
<td>0.103</td>
<td>0.199</td>
<td>0.254</td>
<td>0.254</td>
<td>0.190</td>
<td>0.159</td>
<td>0.231</td>
</tr>
<tr>
<td>Middle-aged</td>
<td>0.177</td>
<td>0.174</td>
<td>0.210</td>
<td>0.225</td>
<td>0.214</td>
<td>0.156</td>
<td>0.210</td>
</tr>
<tr>
<td>Old</td>
<td>0.206</td>
<td>0.183</td>
<td>0.206</td>
<td>0.181</td>
<td>0.224</td>
<td>0.154</td>
<td>0.200</td>
</tr>
<tr>
<td>Total</td>
<td>0.168</td>
<td>0.180</td>
<td>0.218</td>
<td>0.222</td>
<td>0.211</td>
<td>0.155</td>
<td>0.208</td>
</tr>
</tbody>
</table>

3.4.2 Evaluation

This subchapter presents the evaluation results of the proposed MS risk quantification model obtained using six sub-groups. Figure 3.9 and Fig. 3.10 show the regression analysis results of the average ASD values of each sub-group according to the number of MS risk factors exceeding the thresholds of each risk factor. According to the analyzed $R^2$ values, we can judge that there is strong positive correlation be-
tween the ASD values and the number of MS risk factors. Therefore, we can claim that the proposed ASD can effectively represent MS risk.

Figure 3.9: Regression analysis of proposed ASD-based model and the number of MS risk factors for male subjects sub-groups.

Figure 3.10: Regression analysis of proposed ASD-based model and the number of MS risk factors for female subjects sub-groups.

Since the proposed risk quantification model is based on ASD values, it is important to determine clinically evaluated thresholds. Thus, we computed the ASD values of each sub-group using the weights for MS risk factors for the sub-group presented in Table 3.6. Then, we counted the occurrence over ASD values.
In this chapter, the interval of ASD value is set to 0.01. From Fig. 3.11 to Fig. 3.13, we show the analysis results of the number of MS subjects and non-MS subjects over ASD values among male sub-groups. The blue line shows the counts of non-MS subjects over ASD value, whereas the green line shows the counts of MS subjects. The analysis results indicate that the number of MS-present subjects increases as the ASD value increases, which confirms the positive correlation between ASD values and the risk of having MS.

![Figure 3.11](image1.png)

Figure 3.11: Analysis of the numbers of MS subjects and non-MS subjects over ASD value for male sub-groups: young male sub-group. Interval of ASD range=0.01.

![Figure 3.12](image2.png)

Figure 3.12: Analysis of the numbers of MS subjects and non-MS subjects over ASD value for male sub-groups: middle-aged male sub-group. Interval of ASD range=0.01.
To perform in-depth analysis regarding the determination of ASD thresholds, we further divided each sub-group into four cases, as listed in Table 3.8. Among the four cases, Cases 2 and 3 indicate the errors of our proposed model. Since Case 3 subjects are already diagnosed to have MS, we focus on identifying Case 2 subjects when determining ASD thresholds. Furthermore, MS is a chronic disease and requires continuous management, so it is more important to identify potential patients that have high risk for having MS in the future rather than accurate diagnosis. Therefore, we determined the ASD threshold for a sub-group as a lower value of ASD where 50% or more of the subjects are identified as having MS.

Table 3.8: Detailed sub-cases of each sub-group belongs to male subjects.

<table>
<thead>
<tr>
<th>Case 1</th>
<th>A subject whose ASD value exceeds ASD threshold and having MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 2</td>
<td>A subject whose ASD value exceeds ASD threshold and NOT having MS</td>
</tr>
<tr>
<td>Case 3</td>
<td>A subject whose ASD value does NOT exceed ASD threshold and having MS</td>
</tr>
<tr>
<td>Case 4</td>
<td>A subject whose ASD value does NOT exceed ASD threshold and NOT having MS</td>
</tr>
</tbody>
</table>

Figure 3.14 shows in-depth analysis of the numbers of MS subjects and non-MS subjects over ASD value for the young male sub-group. The bar graph in red shows the percentage of MS patients over a given
ASD range. According to the analysis result, when the ASD range is from 0.90 to 0.91 (0.90≤ASD<0.91), the MS patients percentage is 54%. Therefore, the ASD threshold for young male subjects can be determined as 0.90.

Figure 3.14: Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for young male sub-group.

Similarly, the middle-aged male sub-group shows an MS patient percentage of 50%, when the ASD range is from 0.89 to 0.90 (Fig. 3.15). Thus, the ASD threshold for middle-aged male subjects becomes 0.89.

Figure 3.15: Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for middle-aged male sub-group.
For the old male sub-group in Fig. 3.16, the percentage of MS patients first becomes 50% when the ASD range is from 0.61 to 0.62. However, the number of MS subjects and the number of non-MS subjects are 2 and 2, respectively, and the percentage immediately becomes less than 50% after the ASD range. Thus, we have chosen the ASD threshold as the ASD range from 0.81 to 0.82. Therefore, the ASD threshold for old male subjects is 0.81.

Figure 3.16: Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for old male sub-group.

Similar to the results for the male subjects sub-groups, from Fig. 3.17 to Fig. 3.19, we show the analysis results of the number of MS subjects and non-MS subjects over ASD value for the female sub-groups. The blue line shows the counts of non-MS subjects over ASD value, whereas the green line shows the counts of MS subjects. These analysis results also show the correlation between ASD values and risk of having MS in female subjects.
Figure 3.17: Analysis of the numbers of MS subjects and non-MS subjects over ASD value for female subgroups: young sub-group. Interval of ASD range=0.01.

Figure 3.18: Analysis of the numbers of MS subjects and non-MS subjects over ASD value for female subgroups: middle-aged sub-group. Interval of ASD range=0.01.

Figure 3.19: Analysis of the numbers of MS subjects and non-MS subjects over ASD value for female subgroups: old sub-group. Interval of ASD range=0.01.
Figure 3.20 shows in-depth analysis of the numbers of MS subjects and non-MS subjects over ASD value for the young female sub-group. The bar graph in red shows the percentage of MS patients over an ASD range. According to the analysis results, the percentage of MS patients becomes 50% when the ASD range is from 0.83 to 0.84 or from 0.85 to 0.86. However, when the ASD range is from 0.83 to 0.84, the numbers of MS subjects and non-MS subjects are small, and the percentage immediately becomes less than 50% after the ASD range. Therefore, we determined the ASD threshold to be the ASD range from 0.87 to 0.88 (0.87≤ASD<0.88) where the MS patients percentage is 57%. Therefore, we determined the ASD threshold for young female subjects as 0.87.

Figure 3.20: Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for young female sub-group.

Similarly, for the middle-aged female sub-group in Fig. 3.21, we chose the ASD range from 0.85 to 0.86, where the MS patients percentage is 62%. Thus, the ASD threshold for middle-aged female subjects becomes 0.85.
Figure 3.21: Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for middle-aged female sub-group.

For the old female sub-group in Fig. 3.22, we chose the ASD range from 0.82 to 0.83, where the percentage of MS patients is 70%. Therefore, the ASD threshold for the old female sub-group is 0.82. The percentage of MS patients becomes more than 50% when the ASD ranges are from 0.66 to 0.67 or from 0.78 to 0.79. However, the percentage immediately becomes less than 50% after the ASD ranges from 0.66 to 0.67 or from 0.78 to 0.79, so the ASD ranges were not selected.

Figure 3.22: Detailed analysis of the numbers of MS subjects and non-MS subjects over ASD value for old female sub-group.
Table 3.9: Examination results for MS risk factors for case 2 subjects.

<table>
<thead>
<tr>
<th>Sub-groups of Case 2 Subject (ASD threshold)</th>
<th>Number of risk factors (number of subjects)</th>
<th>WC</th>
<th>HDL-C</th>
<th>TG</th>
<th>SBP</th>
<th>DBP</th>
<th>FG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Male (0.90)</td>
<td>0 (3)</td>
<td>82.4±5.3</td>
<td>44.0±2.2</td>
<td>137.7±7.9</td>
<td>122.0±3.6</td>
<td>83.3±0.5</td>
<td>93.3±3.8</td>
</tr>
<tr>
<td></td>
<td>1 (34)</td>
<td>83.8±4.0</td>
<td>42.1±5.4</td>
<td>179.8±43.8</td>
<td>116.8±6.9</td>
<td>78.7±5.8</td>
<td>93.0±10.4</td>
</tr>
<tr>
<td></td>
<td>2 (79)</td>
<td>85.2±6.2</td>
<td>38.0±7.2</td>
<td>216.9±84.1</td>
<td>119.9,10.9</td>
<td>79.8±8.5</td>
<td>92.3±8.3</td>
</tr>
<tr>
<td>Middle-aged Male (0.89)</td>
<td>0 (23)</td>
<td>85.0±3.3</td>
<td>44.0±4.6</td>
<td>128.3±12.8</td>
<td>117.3±6.7</td>
<td>78.3±4.5</td>
<td>91.9±5.8</td>
</tr>
<tr>
<td></td>
<td>1 (130)</td>
<td>83.9±4.5</td>
<td>42.9±8.0</td>
<td>147.9±51.7</td>
<td>119.3±10.8</td>
<td>79.6±6.9</td>
<td>98.1±23.3</td>
</tr>
<tr>
<td></td>
<td>2 (239)</td>
<td>85.9±5.4</td>
<td>40.0±7.6</td>
<td>180.1±89.9</td>
<td>124.4±16.0</td>
<td>83.6±9.9</td>
<td>101.6±28.3</td>
</tr>
<tr>
<td>Old-aged Male (0.81)</td>
<td>0 (5)</td>
<td>81.1±5.5</td>
<td>44.4±3.7</td>
<td>127.2±18.1</td>
<td>116.0±7.0</td>
<td>73.4±7.9</td>
<td>91.6±5.2</td>
</tr>
<tr>
<td></td>
<td>1 (39)</td>
<td>83.7±7.9</td>
<td>45.6±9.9</td>
<td>127.5±32.3</td>
<td>128.9±14.9</td>
<td>77.6±7.3</td>
<td>93.3±6.9</td>
</tr>
<tr>
<td></td>
<td>2 (72)</td>
<td>86.1±7.1</td>
<td>41.7±9.9</td>
<td>134.7±64.9</td>
<td>135.7±20.2</td>
<td>80.1±11.6</td>
<td>100.7±19.5</td>
</tr>
<tr>
<td>Young Female (0.87)</td>
<td>0 (0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1 (2)</td>
<td>82.1±0.4</td>
<td>44.5±8.5</td>
<td>124.5±4.5</td>
<td>146.0±24</td>
<td>96.5±14.5</td>
<td>95.5±1.5</td>
</tr>
<tr>
<td></td>
<td>2 (16)</td>
<td>83.5±6.7</td>
<td>40.3±7.9</td>
<td>195.1±63.5</td>
<td>110.6±5.4</td>
<td>77.1±4.0</td>
<td>94.3±7.4</td>
</tr>
<tr>
<td>Middle-aged Female (0.85)</td>
<td>0 (10)</td>
<td>78.5±3.6</td>
<td>53.6±2.2</td>
<td>129.8±13.0</td>
<td>116.0±9.9</td>
<td>75.4±4.4</td>
<td>92.7±3.3</td>
</tr>
<tr>
<td></td>
<td>1 (81)</td>
<td>79.4±5.1</td>
<td>46.5±8.7</td>
<td>135.4±39.0</td>
<td>117.4±11.7</td>
<td>76.7±5.4</td>
<td>91.4±5.0</td>
</tr>
<tr>
<td></td>
<td>2 (225)</td>
<td>81.9±7.0</td>
<td>42.2±8.4</td>
<td>159.2±87.8</td>
<td>121.3±15.6</td>
<td>79.3±9.1</td>
<td>98.5±24.2</td>
</tr>
<tr>
<td>Old-aged Female (0.82)</td>
<td>0 (19)</td>
<td>75.1±5.9</td>
<td>57.3±4.9</td>
<td>90.6±21.5</td>
<td>118.3±7.7</td>
<td>70.3±7.1</td>
<td>89.0±6.3</td>
</tr>
<tr>
<td></td>
<td>1 (82)</td>
<td>76.7±6.7</td>
<td>49.9±14.5</td>
<td>95.6±30.1</td>
<td>127.3±18.1</td>
<td>77.1±11.1</td>
<td>91.0±7.3</td>
</tr>
<tr>
<td></td>
<td>2 (141)</td>
<td>80.4±8.7</td>
<td>45.0±9.7</td>
<td>119.0±49.1</td>
<td>134.7±19.5</td>
<td>78.2±9.8</td>
<td>97.0±23.1</td>
</tr>
</tbody>
</table>

* Values are means ± SD
Table 3.9 lists the examination results for MS risk factors for Case 2 subjects. According to the analysis results, for young and middle-aged males, triglycerides is the most frequent risk factor for the prevalence of MS, whereas blood pressure, especially systolic blood pressure, is the most frequent factor for old males. For females, HDL cholesterol is the most frequent risk factor for all sub-groups. These analysis results imply that the frequent risk factors occurring in ASD ranges where more than 50% of subjects are MS patients are high triglycerides, low HDL-cholesterol, and high SBP. It is noted that young males and young females show high DBP results with low numbers of risk factors. The main reason for these results is due to the small sample sizes of these sub-groups. Especially, for young female sub-groups having one risk factor, one subject had low HDL-cholesterol value, but the other had hypertension with high SBP and DBP. These caused the high mean and standard deviation of DBP for the young female sub-group.

3.4.3 Discussion

This subchapter discusses several important issues regarding the advantages and disadvantages as well as the applicability of the proposed ASD method.

A. Dependency on the Order of Risk Factors Arranged in the Weighted Radar Chart

As we discussed in the Subchapter 3.3.3, the order of risk factors affects the computed ASD values and the changed ASD values can be calculated based on Corollary 3.1. In this subchapter, we investigate whether the changes of the order of risk factors affect the ASD analysis results including ASD thresholds and the number of sub-cases of each sub-group represented by male and female subjects, respectively. According to Table 3.8, Case 1 and Case 3 indicate the subjects diagnosed as MS, whereas Case 2 and Case 4 refer to the subjects not having MS. Since the main objective of the ASD method is to screen subjects with a high risk of MS, which is difficult to identify using the conventional diagnosis criteria, we focus on the analysis of changes between Case 2 and Case 4. The ASD analysis results presented in the Subchapter 3.4.2 are based on the order of risk factors as shown in Fig. 3.4 (FG, WC, HDL-C, TG, and BP). To investigate the effects of changing the order, we applied an example of changed order of risk factors as FG, WC, TG, HDL-C, and BP to the ASD analysis. Then, the numbers of subjects who had been in Case 2 and re-classified as Case 4 or vice versa
are counted. The changes of ASD thresholds were also computed.

Table 3.10 lists the changes of the relevant numbers. The numbers of subjects who had been classified as Case 2 and were re-classified as Case 4 or vice versa indicate the error of the ASD method. As shown in the table, fewer subjects changed from Case 2 to Case 4 than from Case 4 to Case 2. However, the error percentages of Case 2 subjects are greater than those of Case 4 subjects due to the large difference in the numbers of Case 2 and Case 4 subjects. Since the ASD values of the subjects change as the order changes, the ASD thresholds also change. In our analysis, the new ASD thresholds became lower than the original ASD thresholds. The decrease of ASD thresholds caused more subjects to change from Case 4 to Case 2 than from Case 2 to Case 4. Case 4 indicates subjects whose ASD values are below ASD thresholds and who are not being diagnosed with MS. Thus, as the ASD thresholds decrease, some subjects in Case 4 may exceed the decreased ASD thresholds.

<table>
<thead>
<tr>
<th></th>
<th>Num. of Case 2 subjects</th>
<th>Num. of Case 4 subjects</th>
<th>Num. of Case 2 → Case 4 (%</th>
<th>Num. of Case 4 → Case 2, (%)</th>
<th>ASD_{ORG}</th>
<th>ASD_{NEW}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young</td>
<td>28</td>
<td>600</td>
<td>0 (0)</td>
<td>10 (1.7)</td>
<td>0.90</td>
<td>0.88</td>
</tr>
<tr>
<td>Mid.</td>
<td>71</td>
<td>676</td>
<td>4 (5.6)</td>
<td>19 (2.8)</td>
<td>0.89</td>
<td>0.88</td>
</tr>
<tr>
<td>Old</td>
<td>35</td>
<td>186</td>
<td>1 (2.9)</td>
<td>5 (2.7)</td>
<td>0.81</td>
<td>0.80</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young</td>
<td>13</td>
<td>1048</td>
<td>1 (7.7)</td>
<td>8 (0.8)</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>Mid.</td>
<td>63</td>
<td>1015</td>
<td>3 (4.8)</td>
<td>16 (1.6)</td>
<td>0.85</td>
<td>0.83</td>
</tr>
<tr>
<td>Old</td>
<td>35</td>
<td>229</td>
<td>2 (5.0)</td>
<td>7 (3.1)</td>
<td>0.82</td>
<td>0.81</td>
</tr>
</tbody>
</table>

ASD_{ORG}, threshold with the original order; ASD_{NEW}, threshold with the new order

We have performed an investigation on the effects of changing the order of risk factors on the ASD method and showed that the effects are not significant; however, developing a more formal mathematical model to describe and quantify the effects of changing the order of risk factors is required in for further study.

B. Clinical Benefits and Applicability to Other Populations

According to the evaluation results presented in the previous subchapter, when a subject has an ASD value greater than the threshold and is not determined to have MS based on the conventional MS diagnostic
criteria, we can estimate that the subject has a more than 50% risk of developing MS. Thus, through the proposed risk quantification model that is based on ASD rather than the conventional MS diagnosis methods relying on the number of risk factors, it is possible to estimate the risk of having MS and systematically manage the important risk factors for CVD and T2DM. By doing this, the model proposed in this chapter can contribute to prevention of CVD and T2DM. Also, the proposed risk quantification model that is based on ASD could be applied to the management of other diseases.

Therefore, if a subject whose ASD value exceeds the designated threshold is found to have those MS risk factors, the risk of developing MS, CVD, and T2DM can be reduced by proactive medication and treatment, as well as improvement of living habits.

The proposed model in this chapter determines a reference that comprises the weights and ASD thresholds for MS risk quantification using the medical examination results of a large number of sample subjects. Thus, the change of sampling population may cause the changes of the reference comprising the weights and ASD thresholds. For example, when we applied the proposed method to the recent data obtained from the fourth Korea National Health and Nutrition Examination Survey (KNHANES IV-2008) [52], the ASD threshold values were changed, but the correlation between ASD values and having MS risk was retained. Further, assigning different weights on MS risk factors also reduced the errors of the proposed method, similar to KNHANES III database. Fig. 3.23 and Fig. 3.24 show some of the analysis results using KNHANES IV-2008 database. Fig. 3.23 (a) and (b) show the ASD analysis results of young male subjects of KNHANES IV-2008 database. Fig. 3.23 (a) indicates the analysis result using the reference (weights of MS risk factors and ASD threshold) determined by KNHANES III-2005 database. When the reference of KNHANES III-2005 is used for the analysis, the analysis results also show that the number of MS subjects increases as ASD value increases. However, the ASD threshold determined by KNHANES III-2005 database needs to be updated, because the ASD value where 50% of subjects are MS subjects has been changed. Fig. 3.23 (b) shows the ASD analysis based on the updated weights of risk factors and ASD threshold determined by KNHANES IV-2008 data. We suppose various factors such as lifestyle, nutrition, etc. have affected the changes.
Figure 3.23: Analysis of the numbers of MS subjects and non-MS subjects over ASD value for young male sub-group of KNHANES IV-2008: risk factor weights and ASD threshold of KNHANES III-2005 (a), updated risk factor weights and ASD threshold using KNHANES IV-2008 (b). Interval of ASD range=0.01.

Fig. 3. 24 (a) and (b) show the ASD analysis results of middle-aged female subjects of KNHANES IV-2008 database. The middle-aged female subjects also show the similar characteristics to young male subjects shown in Fig. 3.23.

Table 3.1 lists the weights of MS risk factors and ASD thresholds for young males and females between KNHANES III-2005 and KNHANES IV-2008 database. Since the characteristics of populations chang-
es due to various reasons, the determined weights and ASD threshold values may need to be updated by using periodic health examination results. Further, some countries also provides periodic health examination results for research purpose and the number of countries is expected to be increasing, so our proposed ASD method could would be applicable to other populations in the future.

Table 3.1: Changes of the weights of MS risk factors and ASD thresholds.

<table>
<thead>
<tr>
<th></th>
<th>FG</th>
<th>WC</th>
<th>HDL-C</th>
<th>TG</th>
<th>BP</th>
<th>ASD Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young male (2005)</td>
<td>0.103</td>
<td>0.199</td>
<td>0.254</td>
<td>0.254</td>
<td>0.190</td>
<td>0.90</td>
</tr>
<tr>
<td>Young male (2008)</td>
<td>0.175</td>
<td>0.216</td>
<td>0.164</td>
<td>0.266</td>
<td>0.179</td>
<td>0.91</td>
</tr>
<tr>
<td>Middle-aged female (2005)</td>
<td>0.156</td>
<td>0.210</td>
<td>0.265</td>
<td>0.182</td>
<td>0.187</td>
<td>0.85</td>
</tr>
<tr>
<td>Middle-aged female (2008)</td>
<td>0.206</td>
<td>0.217</td>
<td>0.230</td>
<td>0.196</td>
<td>0.151</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Therefore, the proposed method is adaptively applicable to various populations and the values of ASD thresholds and weights can be computed in accordance with a target population. It is noted that the determined values of weights and ASD threshold may need to be updated by using periodic health examination results because the characteristics of the target population may change over time due to various reasons.

The following summarizes the procedures for utilizing the ASD method.

- **At time** $T_0$: Determine a reference values for MS risk factors’ weights and ASD thresholds using examination data for a population.
- After the determination of the reference, use the established reference for predicting MS risk of individuals.
- **At time** $T_i$: Update the reference using new examination data for the population.
- After the update, use the updated reference for predicting MS risk of individuals.

Fig. 3.25 shows the procedures for applying the proposed ASD method to other populations and database of examination results.
Furthermore, our proposed method can be applied to other populations, if there exist examination results of MS risk factors for large samples. Currently, the US [43], New Zealand [44], and Australia [45] also provide nationwide health examination results for research and public healthcare purposes. Therefore, our proposed method can be applied to those populations. Further, the report published by the US Center for Disease Control and Prevention (CDC) indicated that Brazil, Kuwait, South Africa, Israel, United Kingdom, and France were also planning national health examination surveys and collaborating with CDC [46], so our proposed method could be applied to those populations in the future.

**C. Side-effects and Weaknesses**

The method for deciding the weights of MS risk factors used in this chapter is based on the frequency of each risk factor in the examination results. Therefore, changes of examination subjects cause changes of weights. The selection of KNHANES III sample subjects was performed with consideration of the statistical
characteristics so that the sample subjects would accurately represent the Korean population, but the determination of weights may still be biased, because the examination did not cover the whole Korean population.

Our proposed model quantifies the risk of having MS by using ASD, which is an integrated indicator of MS risk factors. Thus, during the integration procedure based on standardizing the examined values of each risk factor, the originally examined values are integrated with dimensionless values between 0 and 1. This characteristic may cause two subjects having the same ASD value to have different values for each MS risk factor. For example, let us assume that two subjects have the same ASD value of 0.8. Then, it is possible for one subject to have high HDL-C, WC, and TG values and the other subject to have high values of FG, BP, and TG. Thus, it is necessary to draw a weighted radar chart of the subjects’ examination results of MS risk factors to investigate the status of the risk factors in detail. Therefore, the ASD model can be utilized as a screening criterion for subjects with potential risk of MS, and the risk of each MS risk factor can be analyzed by using the weighted radar chart.

D. Comparison with Equally Weighted Method

Since the ASD method allows different weights on MS risk factors, the comparison with the conventional equal weights-based method needs to be discussed in order to quantify the benefit of the ASD method. We performed the comparison using the young male and female subjects. It is important to screen the subjects with high risk of MS (Case 2), which it is difficult to identify using the current MS diagnosis criteria, so we compared the capability for screening Case 2 subjects by counting the number of Case 2 over ASD value. Table 3.12 lists the weights of MS risk factors for comparison between equally weighted ASD method and differently weighted ASD method.

Table 3.12: Weights of MS risk factors for comparison.

<table>
<thead>
<tr>
<th>Sub-group</th>
<th>FG</th>
<th>WC</th>
<th>HDL-C</th>
<th>TG</th>
<th>BP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different weights</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young male</td>
<td>0.103</td>
<td>0.199</td>
<td>0.254</td>
<td>0.254</td>
<td>0.190</td>
</tr>
<tr>
<td>Young female</td>
<td>0.159</td>
<td>0.231</td>
<td>0.297</td>
<td>0.203</td>
<td>0.110</td>
</tr>
<tr>
<td>Equal weights</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young male</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Young female</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
</tbody>
</table>

- 101 -
Fig. 3.26 shows the differences of the number of Case 2 subjects between the ASD method (\(\text{CASE2}_{\text{ASD}}\)) and the equally weighted radar chart method (\(\text{CASE2}_{\text{EWR}}\)) by subtracting the number of \(\text{CASE2}_{\text{EWR}}\) from \(\text{CASE2}_{\text{ASD}}\). As shown in the figure, the conventional equally weighted radar chart method in inferior to identifying Case 2 subjects over ASD value. In the figure, the blue bar and the red bar indicate the ASD thresholds for young male and female subjects determined in the previous Subchapter 3.4.2. By using the number of Case 2 subjects listed in Table 3.9, we can calculate the accuracy for identifying Case 2 subjects using the equally weighted radar chart method as 46.2% for young female subjects and 78.6% for young male subjects, respectively. It is noted that when the ASD threshold is higher than 0.96, the differences of the number of \(\text{CASE2}_{\text{EWR}}\) and \(\text{CASE2}_{\text{ASD}}\) become zero for both male and female subjects. The reason for these is that all subjects whose ASD values are higher than 0.96 are MS diagnosed subjects (Refer to Fig. 3.14 and Fig. 3.20).

![Figure 3.26: Differences of the number of Case 2 between \(\text{CASE2}_{\text{EWR}}\) and \(\text{CASE2}_{\text{ASD}}\) over ASD value.](image)

Fig. 3.27 shows the ratio of total number of Case 2 subjects of \(\text{CASE2}_{\text{EWR}}\) to \(\text{CASE2}_{\text{ASD}}\) over ASD value. In accordance with Fig. 3.26, the numbers of \(\text{CASE2}_{\text{EWR}}\) are smaller than those of \(\text{CASE2}_{\text{ASD}}\) for both subjects.
E. Applicability to Managing Temporal Changes

Since ASD value can serve as an initial criterion for screening risk of having MS, our proposed method helps with management of the temporal changes in MS risk. For example, let us assume that a young male subject regularly examines the values of MS risk factors. At some time, if the ASD value of the subject exceeds the threshold for young males, the statuses of each risk factor will be investigated and risk factors approaching each threshold could be managed by a physician. Also, during the following regular examinations, not only calculation of ASD, but also the generation of a weighted radar chart could be used to analyze the temporal changes of the subject. Once the ASD value of the subject is reduced to below the threshold, only the ASD value would be used to regularly manage the subject’s health.

Furthermore, since patients with chronic diseases including MS typically stay outside of hospitals for the majority of chronic disease care cycles except routine hospital visits, providing application for personal devices can help patients with self-management. Fig. 3.28 shows the prototype implementation of MS risk quantification application. Our prototype system is implemented on the Android 4.2 smartphone platform for easy use. By regularly analyzing the ASD value and the weighted radar chart of MS risk factors, it is possible to manage the temporal progress of MS.
As we mentioned in the Subchapter C, ASD value can serve as an initial criterion for screening risk of having MS. This feature helps with managing the temporal changes in MS risk. For example, as shown in the Fig. 3.29, we can consider the scenario that a young male individual regularly examines the values of MS risk factors. If the ASD value of the person exceeds the threshold of young male at some time, the status of each risk factor will be investigated in detail and risk factors approaching the thresholds will be managed by physician.
Figure 3.29: Temporal changes of ASD value during regular health examinations.

Also, during the following regular examinations, not only calculation of ASD, but also generation of weighted radar chart will be used to analyze the temporal changes of the person (Fig. 3.30). Once the ASD value of the subject becomes below the threshold, only the ASD value will be used to regularly manage the person.

Figure 3.30: Weighted radar chart showing detailed information of MS risk factors.
F. Ancillary Indicator for Distinguishing Subjects Having the Same ASD Value

As we have discussed in the previous subchapter titled “C. Side-effects and Weaknesses”, it is possible for subjects to have the same ASD value because the proposed ASD model maps the normalized examination values of MS risk factors into a single number. To investigate the issue above, we analyzed the occurrence of duplicated ASD values among male and female subjects. The analysis was performed by counting the number of subjects having the same ASD values. Fig. 3.31 shows the distribution of the occurrence of same ASD values for middle-aged male sub-group. The total number of middle-aged male subjects is 1170. The graph shows that most duplicated occurrences occur at ASD value of 1. According to Definition 3.5, the maximum ASD value is 1, so all subjects whose all MS risk factors exceed the corresponding thresholds of risk factors have ASD value of 1. This characteristics cause large number of the occurrences of ASD value 1. Since the number of subjects having ASD value of 1 is much larger than other ASD values, how to differentiate subjects having ASD value 1 is the key issue. Therefore, we develop an ancillary indicator to distinguish subjects having a same ASD value and evaluate the effectiveness of the ancillary indicator for distinguishing the subjects by using the middle ages subjects with ASD value 1. Further, to investigate the effectiveness of distinguishability for subjects having less than ASD value 1, we additionally analyzed the subjects having the second largest occurrences of duplicated ASD value, i.e., the subjects having ASD 0.934 are also analyzed using the proposed ancillary indicator.

![Distribution of the occurrence of same ASD values for middle-aged male sub-group.](image)
Since ASD is calculated by transforming the normalized values of the five MS risk factors into a single number, we develop an ancillary indicator by changing ASD value into number that describes the variation between the risk factors examination values and the corresponding thresholds. To distinguish the duplicated occurrences of ASD values, we develop an ancillary indicator for ASD model as follows.

**Definition 3.6.** Let the sensitivity metric of the MS risk factors ($SM_{i}$) be the weighted distance between the examination values of risk factors and the corresponding thresholds. Then, the sensitivity metric of the MS risk factors is defined as

$$SM_{i} = \sum_{i=1}^{5} (x_{i} - TH_{i})^{2} \times w_{i}$$

(3.17)

, where $x_{i}$ and $TH_{i}$ are the normalized values of $i_{th}$ risk factor examination result and threshold, respectively. $w_{i}$ indicates the weight of $i_{th}$ risk factor.

Eq. (3.17) describes the sensitivity metric between the normalized values of risk factor and the corresponding threshold. The large $SM_{i}$ value implies large difference between examination results and thresholds of MS risk factors. Thus, subjects having the same ASD value can be distinguished by using $SM_{i}$ because those subjects are likely to have different examination results of each risk factor. Fig. 3.32 shows the graphical representation of the parameters in Eq. (3.17).

![Graphical representation of the parameters in Eq. (3.17)](image)

Figure 3.32: Ancillary indicator to distinguish subjects having the same ASD value.
To investigate whether $SM_{rf}$ can distinguish subjects having the same ASD value, we analyzed $SM_{rf}$ value of 49 subjects having ASD 1. Fig. 3.33 shows the distribution of $SM_{rf}$ of 49 middle-aged male subjects having ASD value 1. The graph shows the occurrence of $SM_{rf}$ over $SM_{rf}$ value. Through the analysis, we have observed that three occurrences of same $SM_{rf}$ among subjects, which implies 93.9% capability for distinguishing subjects having a same ASD value.

![Figure 3.33: Distribution of the occurrences of $SM_{rf}$ for the middle-aged male subjects with ASD value 1.](image)

We further analyzed the characteristics of subjects having ASD value 1 and duplicated $SM_{rf}$. Table 3.13 lists the examination results. Since the ASD value of all subjects is 1, the number of MS risk factors exceeding the corresponding thresholds shows value of 5 in accordance with the Definition 3.5. As shown in the table, the number of duplication does not show large occurrences. Also, since all the subjects with duplicated $SM_{rf}$ have all five MS traits, they required detailed investigation by physician.

<table>
<thead>
<tr>
<th>Subject ID</th>
<th>$SM_{rf}$ (%)</th>
<th>WC (cm)</th>
<th>HDL-C (mg/dl)</th>
<th>TG (mg/dl)</th>
<th>SBP (mm Hg)</th>
<th>DBP (mm Hg)</th>
<th>FG (mg/dl)</th>
<th>MS Risk Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>0.84</td>
<td>94.0</td>
<td>38</td>
<td>314</td>
<td>136.00</td>
<td>104.00</td>
<td>102</td>
<td>5</td>
</tr>
<tr>
<td>Subject 2</td>
<td>93.2</td>
<td>93.2</td>
<td>34</td>
<td>294</td>
<td>139.00</td>
<td>90.00</td>
<td>112</td>
<td>5</td>
</tr>
<tr>
<td>Subject 3</td>
<td>0.85</td>
<td>94.2</td>
<td>39</td>
<td>236</td>
<td>126.00</td>
<td>88.00</td>
<td>144</td>
<td>5</td>
</tr>
<tr>
<td>Subject 4</td>
<td>93.2</td>
<td>93.2</td>
<td>36</td>
<td>221</td>
<td>152.00</td>
<td>94.00</td>
<td>114</td>
<td>5</td>
</tr>
<tr>
<td>Subject 5</td>
<td>1.04</td>
<td>90.9</td>
<td>33</td>
<td>262</td>
<td>132.00</td>
<td>85.00</td>
<td>143</td>
<td>5</td>
</tr>
<tr>
<td>Subject 6</td>
<td>92.3</td>
<td>92.3</td>
<td>32</td>
<td>210</td>
<td>155.00</td>
<td>103.00</td>
<td>100</td>
<td>5</td>
</tr>
</tbody>
</table>
We have conducted the same analysis to middle-aged male sub-group subjects with ASD value 0.934, which shows 10 occurrences of subjects. Fig. 3.34 shows the distribution of $SM_{rf}$ of 10 middle-aged male subjects having ASD value 0.934. The graph shows the occurrence of $SM_{rf}$ over $SM_{rf}$ value. Through the analysis, we have observed that all subjects have different $SM_{rf}$ values.

Figure 3.34: Distribution of the occurrences of $SM_{rf}$ for the middle-aged male subjects with ASD value 0.934.

We further analyzed the characteristics of subjects having ASD value 0.934 and duplicated $SM_{rf}$. Table 3.14 lists the examination results. Among the subjects, we further examines whether larger $SM_{rf}$ value precedes higher risk from clinical perspective. For the detailed analysis, we have chosen three subjects, Subject 3, Subject 9, and Subject 10. They have the same ASD value as 0.934 and three MS risk factors, so ASD value-based MS risk quantification model is not able to distinguish the difference between the three subjects. However, the $SM_{rf}$ shows different values, 0.31, 3.14, and 7.23, respectively. The examination results of Subject 3 indicate that all the results are located around the thresholds of each MS risk factor. The examination results of Subject 9 indicate that some of the examination results, particularly TG and FG show large variation on the examination results. TG shows 52% larger than the threshold and FG shows 90% larger than the threshold. According to the clinical diagnosis, large variation on the examination results may indicate higher risk on the health status. Therefore, it is possible to diagnose Subject 9 may have higher risk than Subject 3 according to $SM_{rf}$ value. For Subject 10, the subject also has the same ASD value and number of MS risk factors as Subject 3 or Subject 9. However, the examination results of MS risk factors show that very large value of TG, 326% larger than the corresponding threshold. Thus, the subject is highly likely to have severe problem with TG-related organs or diseases, and it is recommend to immediate consult with a physician. According to the clinical investigation on the three subjects, it is showed that large $SM_{rf}$ implies large difference between the exami-
nation results and thresholds of MS risk factor, which indicates higher possibility to have health problems. Therefore, we can claim that if large difference between examination results and corresponding thresholds indicate higher possibility to have health problem, SM can be used for identifying the subjects having higher health risk among subjects having the same ASD value.

Table 3.1: Examination results of MS risk factors for middle-aged male subjects having ASD value 0.934 and duplicated SM.

<table>
<thead>
<tr>
<th>Subject ID</th>
<th>SM (%)</th>
<th>WC (cm)</th>
<th>HDL-C (mg/dl)</th>
<th>TG (mg/dl)</th>
<th>SBP (mm Hg)</th>
<th>DBP (mm Hg)</th>
<th>FG (mg/dl)</th>
<th>MS Risk Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>0.22</td>
<td>92.1</td>
<td>45</td>
<td>215</td>
<td>132.00</td>
<td>86.00</td>
<td>92</td>
<td>3</td>
</tr>
<tr>
<td>Subject 2</td>
<td>0.24</td>
<td>89.0</td>
<td>32</td>
<td>162</td>
<td>129.00</td>
<td>89.00</td>
<td>91</td>
<td>3</td>
</tr>
<tr>
<td>Subject 3</td>
<td>0.31</td>
<td>81.1</td>
<td>39</td>
<td>186</td>
<td>110.00</td>
<td>80.00</td>
<td>110</td>
<td>3</td>
</tr>
<tr>
<td>Subject 4</td>
<td>0.33</td>
<td>83.2</td>
<td>32</td>
<td>142</td>
<td>112.00</td>
<td>82.00</td>
<td>108</td>
<td>2</td>
</tr>
<tr>
<td>Subject 5</td>
<td>0.37</td>
<td>80.8</td>
<td>38</td>
<td>144</td>
<td>141.00</td>
<td>105.00</td>
<td>104</td>
<td>3</td>
</tr>
<tr>
<td>Subject 6</td>
<td>0.52</td>
<td>98.5</td>
<td>32</td>
<td>124</td>
<td>139.00</td>
<td>95.00</td>
<td>113</td>
<td>4</td>
</tr>
<tr>
<td>Subject 7</td>
<td>1.03</td>
<td>85.5</td>
<td>36</td>
<td>314</td>
<td>142.00</td>
<td>89.00</td>
<td>93</td>
<td>3</td>
</tr>
<tr>
<td>Subject 8</td>
<td>1.10</td>
<td>107.3</td>
<td>42</td>
<td>128</td>
<td>146.00</td>
<td>80.00</td>
<td>111</td>
<td>3</td>
</tr>
<tr>
<td>Subject 9</td>
<td>3.14</td>
<td>87.0</td>
<td>34</td>
<td>229</td>
<td>100.00</td>
<td>70.00</td>
<td>190</td>
<td>3</td>
</tr>
<tr>
<td>Subject 10</td>
<td>7.23</td>
<td>88.7</td>
<td>34</td>
<td>637</td>
<td>100.00</td>
<td>70.00</td>
<td>116</td>
<td>3</td>
</tr>
</tbody>
</table>

We have performed the similar analysis for young female subjects. Fig. 3.35 shows the distribution of the occurrence of same ASD values for young female sub-group. The total number of young female subjects is 1118. The graph shows that most duplicated occurrences occur at ASD value of 0.333. Unlike middle-aged male sub-group, young female sub-group did not show large occurrences of subjects having ASD value 1. As shown in Fig. 3.20, the total number of MS subject for young female sub-group is 57, which is much smaller than other sub-groups. This characteristics cause smaller number of the occurrences of ASD value 1 than middle-aged male sub-group. Further, to investigate the effectiveness of distinguishability for young female sub-group, we analyzed the subjects having the largest occurrences of duplicated ASD value, i.e., the subjects having ASD 0.333. Also, the young female sub-group indicated small occurrences of duplicated ASD value over ASD values.
Figure 3.35: Distribution of the occurrence of same ASD values for young female sub-group

We have conducted the same analysis as middle-aged male sub-group subjects to young female sub-group with ASD value 0.333, which shows 8 occurrences of subjects. Fig. 3.36 shows the distribution of $SM_f$ of 9 young female subjects having ASD value 0.333. The graph shows the occurrence of $SM_f$ over $SM_f$ value. Through the analysis, we have observed that all subjects have different $SM_f$ values. Also, the distribution of $SM_f$ values is ranged from 1% to 3% ranges, which is relatively smaller range than middle-aged male sub-group.

Figure 3.36: Distribution of the occurrences of $SM_f$ for the young female subjects with ASD value 0.333.
We further analyzed the characteristics of subjects having ASD value 0.333 and duplicated $SM_{rf}$. Table 3.15 lists the examination results. Among the subjects, we further examines whether larger $SM_{rf}$ value precedes higher risk from clinical perspective. For the detailed analysis, we have chosen three subjects, Subject 2, Subject 5, and Subject 8. They have the same ASD value as 0.333 and one MS risk factors, so ASD value-based MS risk quantification model is not able to distinguish the difference between the three subjects. However, the $SM_{rf}$ shows different values, 1.81, 2.60, and 3.15, respectively. Since ASD value of the subjects and the number of MS risk factors exceeding the corresponding thresholds are small, the examination results of most MS risk factors indicate less than the corresponding thresholds. Thus, the $SM_{rf}$ values indicate the differences of the examination results below the thresholds. Therefore, the large $SM_{rf}$ value may not necessarily indicate the higher health risk from the perspective of clinical diagnosis.

Table 3.15: Examination results of MS risk factors for young female subjects having ASD value 0.333 and duplicated $SM_{rf}$

<table>
<thead>
<tr>
<th>Subject ID</th>
<th>$SM_{rf}$ (%)</th>
<th>WC (cm)</th>
<th>HDL-C (mg/dl)</th>
<th>TG (mg/dl)</th>
<th>SBP (mm Hg)</th>
<th>DBP (mm Hg)</th>
<th>FG (mg/dl)</th>
<th>MS Risk Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>1.55</td>
<td>69.50</td>
<td>50</td>
<td>58</td>
<td>106.00</td>
<td>70.00</td>
<td>78</td>
<td>0</td>
</tr>
<tr>
<td>Subject 2</td>
<td>1.81</td>
<td>67.00</td>
<td>59</td>
<td>46</td>
<td>112.00</td>
<td>86.00</td>
<td>89</td>
<td>1</td>
</tr>
<tr>
<td>Subject 3</td>
<td>1.89</td>
<td>69.10</td>
<td>61</td>
<td>59</td>
<td>105.00</td>
<td>62.00</td>
<td>89</td>
<td>0</td>
</tr>
<tr>
<td>Subject 4</td>
<td>2.13</td>
<td>71.50</td>
<td>66</td>
<td>84</td>
<td>100.00</td>
<td>58.00</td>
<td>75</td>
<td>0</td>
</tr>
<tr>
<td>Subject 5</td>
<td>2.60</td>
<td>63.40</td>
<td>39</td>
<td>52</td>
<td>112.00</td>
<td>65.00</td>
<td>89</td>
<td>1</td>
</tr>
<tr>
<td>Subject 6</td>
<td>2.80</td>
<td>62.10</td>
<td>42</td>
<td>63</td>
<td>101.00</td>
<td>69.00</td>
<td>83</td>
<td>1</td>
</tr>
<tr>
<td>Subject 7</td>
<td>3.05</td>
<td>61.40</td>
<td>65</td>
<td>75</td>
<td>112.00</td>
<td>79.00</td>
<td>82</td>
<td>0</td>
</tr>
<tr>
<td>Subject 8</td>
<td>3.15</td>
<td>65.30</td>
<td>37</td>
<td>63</td>
<td>91.00</td>
<td>58.00</td>
<td>81</td>
<td>1</td>
</tr>
</tbody>
</table>

In this subchapter, we proposed an ancillary indicator, i.e. sensitivity metric of risk factors, $SM_{rf}$, defined by the average variation of the MS risk factors in order to distinguish subjects having equal ASD values. Through the analyses of the male and female subjects having duplicated ASD values, we showed that the proposed average variation of the MS risk factors can effectively distinguish subjects having equal ASD values. It also investigated that the larger $SM_{rf}$ value may indicate the higher health risk from the clinical diagnosis perspective. Therefore, through the clinical discussion about the proposed $SM_{rf}$, it can be claimed that the $SM_{rf}$ is
effective for further identifying the subjects with higher health risk among subjects having duplicated ASD value.

G. Longitudinal Analysis of ASD Model Using 6-years Examination Results

This subchapter presents retrospective longitudinal analysis results of the proposed ASD using 6-years examination results. A retrospective longitudinal study generally means to take a look back at events that already have taken place. For example, the term is used in medicine, describing a look back at a patient's medical history or lifestyle. To perform the retrospective longitudinal study, we obtained 6-years’ health examination records for volunteers in ETRI (Electronics and Telecommunications Research Institute). The obtained health records consist of the examination results of MS risk factors from 2007 to 2012. We have obtained total nine middle-aged volunteers’ health records (5 males and 4 females). By using the health records, we investigate whether a subject with ASD value over the corresponding threshold actually develops to MS over time period. Table 3.16 shows the characteristics of obtained sampling subjects.

Fig. 3.37 shows the temporal changes of ASD values of sampling subjects of middle-aged male subjects. In the figure, a subject’s ASD value within a shaded rectangular means that the subject was diagnosed as MS at the time frame. The red bar indicates the ASD threshold (0.89) for middle-aged male subjects, which was determined in the Subchapter 3.4.2. In our analysis, all samples exceeding ASD threshold were identified to have MS by using the MS diagnosis criteria. In other words, when a subject’s ASD value exceeds corresponding threshold, the subject is diagnosed as MS in our sample analysis. When a subject’s ASD value is below the corresponding threshold, the subject may be diagnosed as MS in our sample analysis. However, most of the samples below the corresponding ASD threshold were identified not having MS. Therefore, our analysis results show that if a subject’s ASD value exceeds the corresponding threshold and then forthcoming ASD value also exceeds the threshold the subject is highly likely to have MS.
Table 3.16: Characteristics of sampling subjects.

<table>
<thead>
<tr>
<th>Year</th>
<th>Sub-group</th>
<th>WC</th>
<th>HDL-C</th>
<th>TG</th>
<th>SBP</th>
<th>DBP</th>
<th>FG</th>
<th>BMI</th>
<th>ASD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>87.40</td>
<td>42.40</td>
<td>300.00</td>
<td>118.00</td>
<td>78.40</td>
<td>93.60</td>
<td>25.45</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±7.53</td>
<td>±7.92</td>
<td>±328.12</td>
<td>±14.75</td>
<td>±14.50</td>
<td>±13.84</td>
<td>±2.55</td>
<td>±0.22</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>71.75</td>
<td>50.90</td>
<td>103.50</td>
<td>105.00</td>
<td>70.00</td>
<td>89.75</td>
<td>20.89</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±11.69</td>
<td>±9.82</td>
<td>±39.54</td>
<td>±15.00</td>
<td>±7.07</td>
<td>±3.11</td>
<td>±3.41</td>
<td>±0.14</td>
</tr>
<tr>
<td>2008</td>
<td>Male</td>
<td>87.00</td>
<td>46.26</td>
<td>127.80</td>
<td>117.00</td>
<td>78.40</td>
<td>100.00</td>
<td>25.62</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±7.10</td>
<td>±7.60</td>
<td>±40.59</td>
<td>±10.51</td>
<td>±6.97</td>
<td>±24.14</td>
<td>±3.03</td>
<td>±0.20</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>71.00</td>
<td>46.33</td>
<td>97.00</td>
<td>102.50</td>
<td>65.00</td>
<td>91.75</td>
<td>21.07</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±12.21</td>
<td>±12.31</td>
<td>±32.01</td>
<td>±16.39</td>
<td>±8.66</td>
<td>±3.90</td>
<td>±3.31</td>
<td>±0.14</td>
</tr>
<tr>
<td>2009</td>
<td>Male</td>
<td>85.82</td>
<td>44.54</td>
<td>152.80</td>
<td>127.60</td>
<td>81.20</td>
<td>94.20</td>
<td>25.07</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±7.79</td>
<td>±5.24</td>
<td>±84.99</td>
<td>±11.53</td>
<td>±9.06</td>
<td>±22.37</td>
<td>±2.40</td>
<td>±0.23</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>70.48</td>
<td>47.50</td>
<td>98.25</td>
<td>99.00</td>
<td>66.00</td>
<td>92.00</td>
<td>20.93</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±11.24</td>
<td>±11.26</td>
<td>±54.19</td>
<td>±9.11</td>
<td>±6.16</td>
<td>±4.95</td>
<td>±2.64</td>
<td>±0.13</td>
</tr>
<tr>
<td>2010</td>
<td>Male</td>
<td>84.50</td>
<td>43.74</td>
<td>137.40</td>
<td>123.20</td>
<td>79.80</td>
<td>97.80</td>
<td>25.61</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±6.99</td>
<td>±8.81</td>
<td>±29.68</td>
<td>±17.79</td>
<td>±13.56</td>
<td>±26.43</td>
<td>±2.13</td>
<td>±0.14</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>69.85</td>
<td>50.38</td>
<td>101.00</td>
<td>107.25</td>
<td>67.75</td>
<td>90.75</td>
<td>21.07</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±11.34</td>
<td>±4.62</td>
<td>±33.99</td>
<td>±8.67</td>
<td>±10.62</td>
<td>±3.83</td>
<td>±2.49</td>
<td>±0.14</td>
</tr>
<tr>
<td>2011</td>
<td>Male</td>
<td>89.84</td>
<td>41.68</td>
<td>122.60</td>
<td>126.40</td>
<td>78.00</td>
<td>96.80</td>
<td>25.89</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±6.02</td>
<td>±7.25</td>
<td>±28.10</td>
<td>±11.27</td>
<td>±8.02</td>
<td>±17.38</td>
<td>±1.95</td>
<td>±0.12</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>70.33</td>
<td>54.90</td>
<td>116.75</td>
<td>101.75</td>
<td>66.00</td>
<td>89.75</td>
<td>20.39</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±8.11</td>
<td>±4.92</td>
<td>±38.47</td>
<td>±9.71</td>
<td>±7.78</td>
<td>±4.32</td>
<td>±2.50</td>
<td>±0.19</td>
</tr>
<tr>
<td>2012</td>
<td>Male</td>
<td>88.60</td>
<td>45.20</td>
<td>148.20</td>
<td>129.00</td>
<td>77.00</td>
<td>105.20</td>
<td>25.90</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±5.04</td>
<td>±7.76</td>
<td>±49.85</td>
<td>±21.20</td>
<td>±13.74</td>
<td>±34.14</td>
<td>±2.07</td>
<td>±0.15</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>68.68</td>
<td>52.70</td>
<td>69.50</td>
<td>108.00</td>
<td>66.75</td>
<td>93.00</td>
<td>19.71</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±6.01</td>
<td>±7.80</td>
<td>±15.64</td>
<td>±2.74</td>
<td>±8.10</td>
<td>±7.25</td>
<td>±1.81</td>
<td>±0.06</td>
</tr>
</tbody>
</table>

mean±SD
Fig. 3.38 shows the temporal changes of ASD values of sampling subjects of middle-aged female subjects. For female sampling subjects, no subject identified to have MS and all the ASD values of subjects were below the corresponding ASD threshold of middle-aged female sub-group. Thus, it is observed that if a subject’s ASD value is below the corresponding threshold, the subject is highly likely not to have MS, which consistent with the previous analysis of male subjects.
Through the retrospective longitudinal analysis using 6-years health records of middle-aged male and female subjects, we showed that the proposed ASD method is highly correlated with the incidence of MS and is an indicator of the future incidence of MS. Therefore, we can claim that the ASD method can be utilized for identifying potential MS patients and performing proactive management of the identified patients’ health status before the incidence of MS.

**H. Extension of ASD Threshold Determination Method Using Bayesian Model**

In the previous Subchapter 3.4.2, we presented methodology for determining ASD thresholds of male and female subgroups. The thresholds determination methodology is based on the ratio between MS subjects and Non-MS subjects over ASD values. The threshold is determined as the ASD value where the percentage of MS subjects is same as Non-MS subjects. However, the possibility of having MS for subjects whose ASD values are above the corresponding thresholds might need to be presented, so that physicians can utilize the possibility information during diagnosis procedures. This subchapter presents Bayesian model-based possibility model for determining the ASD thresholds.

A Bayesian network model is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network model can be used to compute the probabilities of the presence of various diseases. From the formal point of view, Bayesian networks are directed acyclic graphs whose nodes represent random variables in the Bayesian sense. They may be observable quantities, latent variables, unknown parameters, or hypotheses. Edges represent conditional dependencies and nodes that are not connected represent variables that are conditionally independent of each other. Each node is associated with a probability function that takes as input a particular set of values for the node’s parent variables and gives the probability of the variable represented by the node. For example, if the parents are $m$ Boolean variables then the probability function could be represented by a table of $2^m$ entries, one entry for each of the $2^m$ possible combinations of its parents being true or false. Similar ideas may be applied to undirected, and possibly cyclic, graphs such are called Markov networks [128].

To analyze the probability to have MS, we establish a Bayesian network model for MS risk diagnosis as shown in Fig. 3.39. The proposed ASD model quantifies MS risk and identifies potential patients with high
risk, so the Bayesian network model for MS can be constructed by using a directed acyclic graph from ASD threshold node to MS node.

![Bayesian network model for MS risk diagnosis using ASD thresholds.](image)

Since we categorized the examined subjects into six subgroups according to genders and ages, we can construct six possibility tables. It is noted that the posterior probability for having MS can be calculated by using longitudinal study for subjects whose ASD values exceed the ASD thresholds. However, the longitudinal study requires follow-up monitoring on the subjects for several years, which requires a lot of budget and time, typically several billion KRW and five or more years. Thus, in this dissertation we use KNHANES-III database snapshot-based possibility table [27]. Table 3.17 shows the MS probability table for young male subjects. The table is constructed by using Fig. 3.14 presented in Subchapter 3.4.2.

**Table 3.17: MS probability table for young male subgroup.**

<table>
<thead>
<tr>
<th>Young Male Subgroup (Percent of MS Subjects: 17%)</th>
<th>Exceed ASD Threshold? (ASD≥0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>MS Diagnosed? (Percent of MS Subjects: 17%)</td>
<td></td>
</tr>
<tr>
<td>True</td>
<td>0.06</td>
</tr>
<tr>
<td>False</td>
<td>0.79</td>
</tr>
</tbody>
</table>

According to Table 3.17, ASD threshold value is 76 percent effective in detecting a MS when a young male subject’s ASD value is greater than the ASD threshold of young male subject (0.90). However, the ASD threshold also yields a false positive result for 24 percent of the subjects exceeding the corresponding threshold. That is, if a healthy person is tested, then, with possibility 0.24, the test result will imply he has MS. In our database, 17 percent of the subjects actually have MS, so we can calculate the possibility of a subject has MS given that his ASD value exceeds the ASD threshold for young male subgroup.
Let $D_{YM}$ be the event that the young male subject exceeding the ASD threshold has MS, and $E_{YM}$ the event that he is diagnosed as having MS. The desired probability $P(D_{YM}|E_{YM})$ is obtained by

$$P(D_{YM} \mid E_{YM}) = \frac{P(D_{YM}E_{YM})}{P(E_{YM})} = \frac{P(E_{YM} \mid D_{YM})P(D_{YM})}{P(E_{YM} \mid D_{YM})P(D_{YM}) + P(E_{YM} \mid D_{YM}^C)P(D_{YM}^C)}$$

(3.18)

$$= \frac{(0.76)(0.17)}{(0.76)(0.17) + (0.24)(0.83)}$$

$$= \frac{323}{821} \approx 0.393$$

Therefore, 39.3 percent of those persons whose ASD values exceed the corresponding threshold actually have MS.

We have performed the same analysis methodology to other subgroups of KNHANES-III database. Table 3.18 shows the MS probability table for middle-aged and old male subgroups. The table is also constructed by using ASD analysis results presented in Subchapter 3.4.2.

Table 3.18: MS probability table for middle-aged and old male subgroups.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>MS Diagnosis</th>
<th>Exceed ASD Threshold?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Middle-aged Male Subgroup</td>
<td>MS</td>
<td>True</td>
</tr>
<tr>
<td>(Percent of MS Subjects: 36%)</td>
<td>Diagnosed?</td>
<td>False</td>
</tr>
<tr>
<td>(ASD Threshold = 0.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Male Subgroup</td>
<td>MS</td>
<td>True</td>
</tr>
<tr>
<td>(Percent of MS Subjects: 36%)</td>
<td>Diagnosed?</td>
<td>False</td>
</tr>
<tr>
<td>(ASD Threshold = 0.81)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to Table 3.18, ASD threshold value for middle-aged male subgroup is 82 percent effective in detecting a MS when a middle-aged male subject’s ASD value is greater than the ASD threshold of middle-aged male subgroup (0.89). However, the ASD threshold for middle-aged male subgroup also yields a false positive result for 18 percent of the subjects exceeding the corresponding threshold. In our database, 36 percent of the middle-aged male subgroup actually has MS, so we can calculate the possibility of a subject has MS given that his ASD value exceeds the ASD threshold for middle-aged male subgroup.
Let $D_{MM}$ be the event that the middle-aged male subject exceeding the ASD threshold has MS, and $E_{MM}$ the event that he is diagnosed as having MS. The desired probability $P(D_{MM}|E_{MM})$ is obtained by

$$P(D_{MM} | E_{MM}) = \frac{P(D_{MM} \cap E_{MM})}{P(E_{MM})} = \frac{P(E_{MM} | D_{MM})P(D_{MM})}{P(E_{MM} | D_{MM})P(D_{MM}) + P(E_{MM} | D_{MM}^c)P(D_{MM}^c)} \quad (3.19)$$

$$= \frac{(0.82)(0.36)}{(0.82)(0.36) + (0.18)(0.64)} \approx 0.719$$

Similarly, ASD threshold value for old male subgroup is 76 percent effective in detecting a MS when an old male subject’s ASD value is greater than the ASD threshold of old male subgroup (0.81). However, the ASD threshold for old male subgroup also yields a false positive result for 24 percent of the subjects exceeding the corresponding threshold. In our database, 36 percent of the old male subgroup actually has MS, so we can calculate the possibility of a subject has MS given that his ASD value exceeds the ASD threshold for old male subgroup.

Let $D_{OM}$ be the event that the middle-aged male subject exceeding the ASD threshold has MS, and $E_{OM}$ the event that he is diagnosed as having MS. The desired probability $P(D_{OM}|E_{OM})$ is obtained by

$$P(D_{OM} | E_{OM}) = \frac{P(D_{OM} \cap E_{OM})}{P(E_{OM})} = \frac{P(E_{OM} | D_{OM})P(D_{OM})}{P(E_{OM} | D_{OM})P(D_{OM}) + P(E_{OM} | D_{OM}^c)P(D_{OM}^c)} \quad (3.20)$$

$$= \frac{(0.76)(0.36)}{(0.76)(0.36) + (0.24)(0.64)} \approx 0.640$$

Similar to male subgroups, we constructed MS probability table for young, middle-aged, and old female subgroup, respectively. Table 3.19 shows the probability table for each subgroup.
Table 3.19: MS probability table for young, middle-aged, and old female subgroups.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>MS Diagnosis</th>
<th>Exceed ASD Threshold?</th>
<th>Effectiveness for Detecting MS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
</tr>
<tr>
<td>Young Female Subgroup</td>
<td>MS</td>
<td>True</td>
<td>0.02</td>
</tr>
<tr>
<td>(Percent of MS Subjects: 5%)</td>
<td>Dianaed?</td>
<td>False</td>
<td>0.94</td>
</tr>
<tr>
<td>(ASD Threshold = 0.87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-aged Female Subgroup</td>
<td>MS</td>
<td>True</td>
<td>0.07</td>
</tr>
<tr>
<td>(Percent of MS Subjects: 26%)</td>
<td>Dianaed?</td>
<td>False</td>
<td>0.70</td>
</tr>
<tr>
<td>(ASD Threshold = 0.85)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Female Subgroup</td>
<td>MS</td>
<td>True</td>
<td>0.10</td>
</tr>
<tr>
<td>(Percent of MS Subjects: 48%)</td>
<td>Dianaed?</td>
<td>False</td>
<td>0.45</td>
</tr>
<tr>
<td>(ASD Threshold = 0.82)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

By utilizing Eq. (3.18), we can calculate the possibility of a subject has MS given that his ASD value exceeds the ASD threshold for young, middle-aged, and old female subgroups as shown in Table 3.20.

Table 3.20: Possibility of a subject has MS given that the ASD value exceeds the corresponding thresholds for young, middle-aged, and old female subgroups.

<table>
<thead>
<tr>
<th>Female Subgroup</th>
<th>Possibility for having MS when the ASD value exceeds the corresponding threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Female</td>
<td>0.119</td>
</tr>
<tr>
<td>Middle-aged Female</td>
<td>0.600</td>
</tr>
<tr>
<td>Old Female</td>
<td>0.840</td>
</tr>
</tbody>
</table>

In this subchapter, we have analyzed the possibility for having MS when a subject’s ASD value exceeds the corresponding threshold value. To investigate the possibility, we have used Bayesian network model in order to describe the relationship between ASD values and MS.
3.5 Summary

To resolve the weaknesses of current MS diagnosis methods, this chapter proposed a risk quantification model for MS, which is based on areal similarity degree analysis between weighted radar charts consisting of MS diagnostic criteria and examination results of MS risk factors of a subject. The clinical effectiveness of the proposed model was evaluated using data of a large number of subjects obtained from the KNHANES III. The evaluation results showed that the proposed model can quantify the risk of MS and effectively identify a group of subjects who can be classified into a potential risk group for having MS in the future. Using the proposed model, we can identify potential MS patients early and monitor the temporal change of the patients’ statuses.

This chapter contributed to identifying potential MS subjects with high risk by using the novel ASD analysis, where the definitions and diagnostic criteria for MS are ambiguous. Further, it presented the risk quantification model that can be used to prevent the incidence of MS, CVD, and T2DM through managing the status of risk factors of the identified potential MS subjects. The limitations and recommendations for further study can be summarized as follows:

- The proposed method quantifies the risk of MS based on ASD values and determines ASD thresholds for identifying individuals with MS risk. However, to further test the validity of the ASD method, a longitudinal study should be performed to determine ASD values in a cohort of subjects at one time point, and investigate whether the ASD values are predictive of the subsequent development of MS. Although we have performed a retrospective longitudinal study using 6-years health records in chapter, the size of sampling group was very limited. Thus, more large scale study would be necessary for more close analysis.

- To apply the proposed method to other populations, examination results of sample subjects that statistically represent the populations are required. Without the collected examination data, it is difficult to use this method.

- The relative weights of MS risk factors have been determined by counting the frequency of each risk factor. However, to accurately determine the relative weights, further study should be performed. Several previous studies investigating the relative importance among MS risk factors can be found in [5][47][49-50].
Chapter 4. Temporal Chronic Disease Progress Model for Clinical Decision Support System

This chapter 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 ASD values between consecutive examination results of MS risk factors. We evaluate the clinical effectiveness of the temporal progress model by using sample subjects’ examination results that has 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.

4.1 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. There have been efforts to develop a metabolic syndrome 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 [56]. However, the development of methods for temporal progress management of metabolic syndrome has not been widely investigated. Different from acute disease, chronic disease requires long term care and its temporal information plays 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 chapter, 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 novel similarity degree-based chronological distance analysis method [56][63]. The proposed method can be used to advanced chronic disease care services in which it combines the chronic disease diagnosis services and temporal progress care services. After that, to efficiently manage patients with MS, we propose a patient status classification method for MS care based on the quantified changes of disease status. We evaluate the proposed patient status classification method using data obtained from the third Korea National Health and Nutrition Examination Survey (KNHANES III) among non-institutionalized civilians in the Republic of Korea, which was conducted by the Korean Ministry of Health and Welfare in 2005 [27]. As we discussed in Chapter 1, an individual having chronic diseases 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 method to be utilized for investigating the at-home management of chronic disease. Finally, we summarize the contributions of this chapter and list the limitations of proposed method and further study issues.
4.2 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 [20]. 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 [53-55].

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. One of mostly popular systems for medical services is health information management system (HIMS). 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. The traditional HIMS 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 [60].
4.3 Model for Temporal Progress of MS Using Chronological Clustering

4.3.1 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 fusion 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 $\alpha$, is the clustering tuning indicator. This indicator is the primary element in defining the similarity between adjacent observations. The indicator $\alpha$ is the tolerance ratio of similarity or correlation between observations [57-58].

The chronological clustering method has been applied to investigate the temporal progress of chronic diseases [59]. 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, Hazzemi et al. [59] developed a distance function and a sensitivity of change function as follows:

**Definition 4.1.** 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 4.2.** Given two consecutive medical report values containing $N$ examination criteria, the distance function $d_N(2)$ is defined as:

$$d_N(2) = \sum_{i=1}^{N} \frac{|\text{Previous normalized report value} - \text{Current normalized report value}|}{\text{Number of examination criteria}}.$$ (4.1)
Let the previous normalized report value and the current normalized report value be $VP_i$ and $VC_i$, respectively and $N$ denotes 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

$$d_N(2) = \sum_{i=0}^{2} |VC_i - VP_i| - \frac{1}{2}.$$

(4.2)

**Definition 4.3.** Given the sensitivity level denoted by $\alpha$ and distance function $d_N$, the importance of information (IoF) for two normalized medical reports containing $N$ examination criteria each is defined as:

$$IoF = \begin{cases} 0, & \text{if } d_N(2) \leq 1 - \alpha \quad \text{NOT noticeable change} \\ \text{yields} \quad \text{Discard it} \end{cases}$$

$$1, & \text{if } d_N(2) > 1 - \alpha \quad \text{Noticeable change} \\ \text{yields} \quad \text{Visualize it} \end{cases}$$

(4.3)

The value of $d_N(2)$ in Eq. (4.2) is related to the indicator $\alpha$ 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 $\alpha$, it is important to decide the value of $\alpha$ that is clinically meaningful. The range of $\alpha$ is from 0 to 1, where the correlation between the current and previous report becomes weaker as $\alpha$ becomes closer to 0. The correlation becomes stronger as $\alpha$ becomes closer to 1. A weak correlation, that is, when $\alpha$ 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 $\alpha$ is 0.1, then the reports are considered similar whenever there is more than 10% similarity between the two reports. Also, if $\alpha$ is 0.8, then the reports considered to be similar whenever there is more than 80% similarity between them [59]. Therefore, the effectiveness of the method is largely dependent on the determination of $\alpha$.

However, the authors of [59] did not provide any further information on determining the value of $\alpha$ 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 areal similarity degree (ASD) based chronological clustering method to identify the temporal progress of chronic disease, especially metabolic syndrome. 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. [56][25]. The authors 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 [2]:

- Equally weighted MS diagnostic criteria
- No explanation for disease status and changes of status
- No information about risk quantification by comparing examination results with diagnostic criteria

Therefore, 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. 4.1.

**Definition 4.4.** According to [56], the ASD of a partial weighted radar chart comprising two polygons $A_{ij}$, i.e., $\Delta OA_{i}A_{j}$, and $B_{ij}$, i.e., $\Delta OB_{i}B_{j}$, is defined as

$$ASD = S(A_{ij} \mid B_{ij}) = \frac{\text{Area of intersection of polygon } A_{ij} \text{ and } B_{ij}}{\text{Area of polygon } B_{ij}}$$ (4.4)

Figure 4.1: A partial radar chart consisting of two polygons. The figure has been adopted from [56] and modified.
Theorem 4.1. 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:

\[
d_2(2) = 2 - \text{ASD}(A_j | B_j) - \text{ASD}(B_j | A_j),
\]

(4.5)

where \(A_j\) and \(B_j\) are partial weighted radar charts constructed using the normalized values of two medical examination criteria, namely \(i\) and \(j\) in each medical report and \(j = (i + 1) \mod 2\).

Proof.

Let us assume that two medical reports \(A_j\) and \(B_j\) are results of the examination criteria of the current and previous medical reports expressed by a partial weighted radar chart. Then, according to Eq. (4.2), \(d_2(2)\) becomes

\[
d_2(2) = \sum_{i=1}^{2} \left| A_{ij} - B_{ij} \right| / 2
\]

(4.6)

\[
= \left| A_{1.2} - B_{1.2} \right| + \left| A_{2.1} - B_{2.1} \right| / 2
\]

\[
= \left| A_{ij} - B_{ij} \right| + \left| A_{ij} - B_{ij} \right| / 2
\]

According to Fig. 4.1, \(A_{ij} - B_{ij}\) equals to \(A_{ij} - B_{ij}\). Thus, Eq. (4.6) becomes \(A_{ij} - B_{ij}\). Since \(A_{ij} - B_{ij}\) indicates the normalized difference between \(A_j\) and \(B_j\), and \(A_j\) and \(B_j\) can be mapped into a scalar value by means of their areas. Therefore, it becomes

\[
\left| A_{ij} - B_{ij} \right| = \frac{\text{Area of } A_j \cdot \text{Intersection of } A_j \cap B_j}{\text{Area of } A_j} + \frac{\text{Area of } B_j \cdot \text{Intersection of } A_j \cap B_j}{\text{Area of } B_j}.
\]

(4.7)

According to Definition 4.4, since the ASD of a partial radar chart is defined as the ratio of the intersection of \(A_j\) and \(B_j\) over \(B_j\), Eq. (4.7) becomes

\[
\left| A_{ij} - B_{ij} \right| = 1 - s(A_j | B_j) + 1 - s(B_j | A_j)
\]

\[
= 2 - s(A_j | B_j) - s(B_j | A_j)
\]

Therefore, we can obtain

\[
d_2(2) = 2 - \text{ASD}(A_j | B_j) - \text{ASD}(B_j | A_j).
\]

(4.8)

(q.e.d)
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 4.1 makes it easy to test whether the examined subject’s disease status has changed over the patient specific sensitivity level described in Definitions 4.2 and 4.3. 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 [56]. Therefore, Theorem 4.1 can be utilized as a preliminary diagnosis method in clinical decision support system for chronic disease.

**Corollary 4.1.** Let \( d_N(2) \) be the distance of two chronologically ordered medical reports containing \( N \) examination criteria in each. According to Theorem 4.1, \( d_N(2) \) can be calculated as:

\[
d_N(2) = 2 - \text{ASD}(R_A | R_B) - \text{ASD}(R_B | R_A)
\]

\[\text{(4.9)}\]

where \( R_A \) and \( R_B \) are weighted radar charts constructed using \( N \) disjoint partial weighted radar charts, \( R_A = \{A_{12}, A_{23}, \ldots, A_{n3}\} \) and \( R_B = \{B_{12}, B_{23}, \ldots, B_{n3}\} \). 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 4.1 provides the calculation of distance function for two partial weighted radar charts, Corollary 4.1 is obvious.

According to Eq. (4.9), we can calculate the distance function \( d_N(2) \) for two weighted radar charts containing 5 examination criteria shown in Fig. 4.2 as 0.31.
Figure 4.2: Two weighted radar charts constructed from two data sets A and B. The figure has been adopted from [56] and modified.

4.3.2 Evaluation and Discussion

This subchapter presents a method to determine a patient’s temporal disease progress based on chronological distance described in Corollary 4.1. Since it had been investigated that ASD can be used to quantifying and identifying risk of MS [56], we can assume that the variation of ASD over time, which is defined as a chronological distance in this chapter, 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 4.1. 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:

$$ y(T) = aT + b $$

(4.10)

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 4.5.** Chronological distance variance over $t_i$ and $t_{i+1}$ is defined as

$$ |y(t_i) - y'(t_{i+1})| \cdot \delta $$

(4.11)

Let $\delta$ be the acceptable range of variation, then we claim that the chronological distance variance is within an control range if

$$ |y(t_i) - y'(t_{i+1})| < y(t_{i+1}) \cdot \delta $$

(4.12)

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. (4.10) and Eq. (4.11), we can obtain the following observations.

In Eq. (4.9), the slope $a$ implies the temporal trend of subject’s disease status. The values of $a$ is defined as follows:

- $a=0$ 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. (4.11), the chronological distance variance quantifies the integrated change of examination results for disease criteria.

$$ |y(t_i) - y'(t_{i+1})| < y(t_{i+1}) \cdot \delta $$

implies the subject’s disease change is within a control range.

$$ |y(t_i) - y'(t_{i+1})| \geq y(t_{i+1}) \cdot \delta $$

implies the subject’s disease change is outside a control range.

To verify the proposed Hypothesis 4.1, we performed an evaluation with a sample patient’s medical examination results. Table 4.1 shows the characteristics of the study subject over time.
Table 4.1: A sample patient’s medical examination results for metabolic syndrome risk factors.

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

BMI, body mass index; BP, blood pressure

Then, we obtained the linear regression result as shown in Fig. 4.3. 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.

![Figure 4.3: Linear regression of ASD values over time.](image)

The calculated chronological distances using Table 4.1 and Eq. (4.9) are shown in Table 4.2.
Table 4.2: Calculation of chronological distance of a sample patient.

<table>
<thead>
<tr>
<th>Time</th>
<th>t₀-t₁</th>
<th>t₁-t₂</th>
<th>t₂-t₃</th>
<th>t₃-t₄</th>
<th>t₄-t₅</th>
<th>t₅-t₆</th>
<th>t₆-t₇</th>
<th>t₇-t₈</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronological distance</td>
<td>0.042</td>
<td>0.064</td>
<td>0.021</td>
<td>0.110</td>
<td>0.034</td>
<td>0.001</td>
<td>0.062</td>
<td>0.063</td>
</tr>
</tbody>
</table>

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

\[
0.2946 < A(t_{i+1}) < 0.4910
\]  

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. (4.12), the subject is likely to have major change at disease status and needs to consult physician. The analysis in this subchapter 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 \( \delta \). Also, longitudinal study using cohort is need to extensively evaluate the clinical effectiveness of the proposed temporal progress identification method.

### 4.3.3 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 subchapter, 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. (4.12). 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. The estimation can be done by using linear regression method. To analyze the effectiveness of linear regression of ASD values in estimating future ASD values, we performed longitudinal analysis using two sample subjects. Table 4.3 and 4.4 show the medical examination results of the two subjects.
Table 4.3: Medical examination results of Subject 1 for retrospective clinical analysis (Young male).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, Years</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
</tr>
<tr>
<td>FG (mg/dl)</td>
<td>92</td>
<td>91</td>
<td>96</td>
<td>89</td>
<td>89</td>
<td>87</td>
<td>94</td>
<td>85</td>
<td>91</td>
</tr>
<tr>
<td>WC (cm)</td>
<td>80</td>
<td>77</td>
<td>82</td>
<td>83</td>
<td>78</td>
<td>77</td>
<td>79</td>
<td>82</td>
<td>77</td>
</tr>
<tr>
<td>HDL-C (mg/dl)</td>
<td>58</td>
<td>49</td>
<td>61</td>
<td>47</td>
<td>58</td>
<td>42</td>
<td>61</td>
<td>63</td>
<td>57</td>
</tr>
<tr>
<td>TG (mg/dl)</td>
<td>120</td>
<td>97</td>
<td>84</td>
<td>91</td>
<td>70</td>
<td>87</td>
<td>83</td>
<td>78</td>
<td>94</td>
</tr>
<tr>
<td>SBP (mmHg)</td>
<td>116</td>
<td>120</td>
<td>110</td>
<td>110</td>
<td>100</td>
<td>120</td>
<td>110</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>DBP (mmHg)</td>
<td>64</td>
<td>80</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>60</td>
<td>60</td>
<td>70</td>
<td>75</td>
</tr>
</tbody>
</table>

Table 4.4: Medical examination results of Subject 2 for retrospective clinical analysis (Middle-aged male).

<table>
<thead>
<tr>
<th>Exam. Time</th>
<th>T1 ('03)</th>
<th>T2 ('04)</th>
<th>T3 ('05)</th>
<th>T4 ('06)</th>
<th>T5 ('07)</th>
<th>T6 ('08)</th>
<th>T7 ('09)</th>
<th>T8 ('10)</th>
<th>T9 ('11)</th>
<th>T10 ('12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, Years</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
</tr>
<tr>
<td>FG (mg/dl)</td>
<td>99</td>
<td>108</td>
<td>105</td>
<td>130</td>
<td>116</td>
<td>147</td>
<td>138</td>
<td>148</td>
<td>129</td>
<td>173</td>
</tr>
<tr>
<td>WC (cm)</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>94</td>
<td>96</td>
<td>98</td>
<td>94</td>
<td>97</td>
<td>88</td>
</tr>
<tr>
<td>HDL-C (mg/dl)</td>
<td>50</td>
<td>61</td>
<td>42.4</td>
<td>53.4</td>
<td>35.5</td>
<td>50.4</td>
<td>46.5</td>
<td>51.7</td>
<td>53.6</td>
<td>54.4</td>
</tr>
<tr>
<td>TG (mg/dl)</td>
<td>190</td>
<td>137</td>
<td>216</td>
<td>153</td>
<td>250</td>
<td>133</td>
<td>206</td>
<td>125</td>
<td>101</td>
<td>172</td>
</tr>
<tr>
<td>SBP (mmHg)</td>
<td>143</td>
<td>140</td>
<td>110</td>
<td>140</td>
<td>130</td>
<td>124</td>
<td>130</td>
<td>125</td>
<td>120</td>
<td>122</td>
</tr>
<tr>
<td>DBP (mmHg)</td>
<td>94</td>
<td>100</td>
<td>90</td>
<td>90</td>
<td>100</td>
<td>82</td>
<td>90</td>
<td>90</td>
<td>85</td>
<td>76</td>
</tr>
</tbody>
</table>
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. 4.4 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

\[ y(t) = -0.0238x(t) + 0.5602, \quad R = 0.7812. \]  

(4.13)

Using Eq. (4.13), we can estimate ASD value at time 2012 as 0.322. The examined ASD value at time 2012 by using Table 4.3 can be calculated as 0.471. Thus, the error of estimation is calculated as 31.57%.

![Figure 4.4: Linear regression of ASD values of Subject 1.](image)

Similarly, we performed linear regression for Subject 2’s ASD values over time and Fig. 4.5 shows the regression results. The linear function of Subject 2 is as

\[ y(t) = -0.0139x(t) + 0.9859, \quad R = 0.4771. \]  

(4.14)

Using Eq. (4.14), we can estimate ASD value at time 2012 as 0.847. The examined ASD value at time 2012 by using Table 4.4 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.

$$y_{WC}(t_i) = a_{WC}x(t_i) + b_{WC}$$
$$y_{HC}(t_i) = a_{HC}x(t_i) + b_{HC}$$
$$y_{TG}(t_i) = a_{TG}x(t_i) + b_{TG}$$
$$y_{BP}(t_i) = a_{BP}x(t_i) + b_{BP}$$
$$y_{FG}(t_i) = a_{FG}x(t_i) + b_{FG}.$$  (4.15)

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. (4.15). Fig. 4.6 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 $t_i$.

Figure 4.5: Linear regression of ASD values of Subject 2.

![Graph showing linear regression of ASD values over time](image)

$y = -0.0139x + 0.9859$

$R^2 = 0.2277$
Regression of MS risk factors

Figure 4.6: Application of linear regressions of MS risk factors to future ASD estimation.

Fig. 4.7 and Fig. 4.8 show the liner 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.

Figure 4.7 Liner regression results for MS risk factors of Subject 1.
Table 4.5: Summary of ASD estimation accuracy comparison.

<table>
<thead>
<tr>
<th>Estimated ASD Value at Time $t_i$ (2012)</th>
<th>Subject</th>
<th>ASD-based Regression</th>
<th>Risk Factors-based Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>0.322</td>
<td>0.350</td>
<td></td>
</tr>
<tr>
<td>Subject 2</td>
<td>0.847</td>
<td>0.896</td>
<td></td>
</tr>
<tr>
<td>Examined ASD Value at Time $t_i$ (2012)</td>
<td>Subject 1</td>
<td>0.471</td>
<td></td>
</tr>
<tr>
<td>Subject 2</td>
<td>0.902</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimation Error (%)</td>
<td>Subject 1</td>
<td>31.57</td>
<td>25.72</td>
</tr>
<tr>
<td>Subject 2</td>
<td>6.10</td>
<td>0.69</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.5 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.

4.3.4 Summary

In this subchapter, we proposed a method for identifying the temporal progress of metabolic syndrome using chronological clustering. 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 metabolic syndrome. The clinical effectiveness of the proposed model was evaluated using a sample patient data and the evaluation results showed that the proposed model can be used to analyze the temporal progress of chronic disease, especially the risk of metabolic syndrome. Using the proposed model, we can effectively manage the patients having metabolic syndrome using the control range of patient status.
4.4 MS Status Tier Classification for Patients Using Control Range of Chronological Clustering

4.4.1 Determination of Control Range Thresholds for Patient Tier Classification based on Chronological Clustering

This subchapter describes the determination of patient tier classification thresholds based on the sensitivity level of the MS risk quantification based on the areal similarity degree (ASD) model that is defined by Theorem 3.2. The theorem quantifies the risk of having MS in the future by analyzing the ASD between health examination results of a patient and the defined thresholds for MS. The clinical effectiveness of the theorem was evaluated using sample examination data collected by a large number of sample subjects. Then, Theorem 4.1 describes the temporal changes of a patient’s MS status by using the chronological distance of two consecutive ASD values over time. In Eq. (4.12), it is determined that whether the MS status changes are within the control ranges of the patient by comparing the chronological distance variance $\delta$ with the difference of the patient’s MS risk factors examination results between two consecutive time. If the difference between the two consecutive examination results is less than $\delta$, it is determined that the patient’s disease status is under control. Therefore, to apply the Theorem 4.1 to identify the temporal changes of a patient’s MS status, it is necessary to develop reference values for $\delta$. In this subchapter, we determine the quantitative values for the references of $\delta$.

Fig. 4.9 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 graph indicate that there are minor, major, and significant change between the current and the previous reports, 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 $\alpha$ 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 current report value and the previous one 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 $\alpha$ 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 $\alpha$ is determined as $TH_j$ to indicate changes including minor change of patient disease status. We chose 0.682 ($1\sigma$) for the value of $TH_j$. 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.

![Diagram showing sensitivity for temporal change of MS status according to different values of $\delta$.](image)

Figure 4.9: Sensitivity for temporal change of MS status according to different values of $\delta$.

According to the disease care service levels of metabolic syndrome shown in Fig. 4.9, 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 4.6 lists the criteria for patients tier classification. We consider MS only in this subchapter, so the Table 4.6 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.

Table 4.6: Patient’s TIER classification based on $d_N(2)$ thresholds.

<table>
<thead>
<tr>
<th>Patient TIER (Threshold for chronological distance variance)</th>
<th>TIER(1)</th>
<th>TIER(2)</th>
<th>TIER(3)</th>
<th>TIER(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Health Risk</td>
<td>Very Low</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
</tbody>
</table>

To establish the values of the thresholds for tiers of patients MS status classification, we performed MS patients status classification analysis using a large number of clinical data. The analysis is based on data obtained from the third Korea National Health and Nutrition Examination Survey (KNHANES III) among non-institutionalized civilians in the Republic of Korea, which was conducted by the Korean Ministry of Health and Welfare in 2005. This survey was a nationwide representative study using a stratified, multistage probability sampling design for the selection of household units. The survey consisted of the following 4 components: the Health Interview Survey, the Health Behavior Survey, the Health Examination Survey, and the Nutrition Survey [27]. A total of 34,145 individuals from these sampling frames were included in the health interview survey; among them, 25,161 subjects aged over 20 years were identified as potential participants in our study. We excluded those with incomplete data for the standardized analysis. This resulted in a final analytical sample of 5,355 subjects (2276 male, 3079 female), aged over 20 years. A total of 5,355 subjects (aged over 20 years) were included in this chapter. The proportion of female subjects was higher than male (57.47% vs. 42.53%). The mean ages of the male and female subjects were 47.22±14.61 and 46.99±15.62 years, respectively. The average BMI was 23.99±3.10 and 23.52±3.38 kg/m2 for male and female subjects, respectively. The percentage of subjects with diabetes mellitus in the male group was higher than that in the female group (6.90% vs. 4.48%). Also, 22.98% of the male subjects had hypertension, whereas
14.68% of the female subjects did. In this chapter, we have classified total subjects into two subject groups by
gender and further classified each subject group into three sub-groups by age: young-adult (from 20 to 39
years old), middle-aged (from 40 to 64 years old), and old-aged (more than 65 years old), respectively.
Therefore, we use a total of six sub-groups for the evaluation of our proposed risk quantification model [56].

To perform in-depth analysis regarding the determination of the thresholds of tiers of the patients MS
status classification, we further divided each sub-group into four Cases, as listed in Table 4.7. Since the
objective of the patients MS status classification is to categorize MS patients according to temporal status
changes, we chose thresholds of tiers for the patients MS status classification based on the incidence of MS
disease, i.e., the percentages of patients over given ASD values. Table 4.8 shows the criteria for determining
MS status changes thresholds.

Table 4.7: Detailed sub-cases of each subject-group [56].

<table>
<thead>
<tr>
<th>Case</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>A subject whose ASD value exceeds ASD threshold and having MS</td>
</tr>
<tr>
<td>Case 2</td>
<td>A subject whose ASD value exceeds ASD threshold and NOT having MS</td>
</tr>
<tr>
<td>Case 3</td>
<td>A subject whose ASD value does NOT exceed ASD threshold and having MS</td>
</tr>
<tr>
<td>Case 4</td>
<td>A subject whose ASD value does NOT exceed ASD threshold and NOT having MS</td>
</tr>
</tbody>
</table>

Table 4.8: Criteria for determining thresholds of patients MS status changes.

<table>
<thead>
<tr>
<th>Patient TIER</th>
<th>TIER(1)</th>
<th>TIER(2)</th>
<th>TIER(3)</th>
<th>TIER(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$(d_n(2) &lt; TH_1)$</td>
<td>$(TH_1 \leq d_n(2) &lt; TH_2)$</td>
<td>$(TH_2 \leq d_n(2) &lt; TH_3)$</td>
<td>$(TH_3 \leq d_n(2))$</td>
</tr>
<tr>
<td>Criteria for</td>
<td>Less than 34.1%</td>
<td>Less than 47.7%</td>
<td>Less than 49.8%</td>
<td>Greater than 49.8%</td>
</tr>
<tr>
<td>TIER Thresholds</td>
<td>(1σ)</td>
<td>(2σ)</td>
<td>(3σ)</td>
<td></td>
</tr>
</tbody>
</table>

The previous criteria are based on the control ranges of chronological distance variance of ASD
values, of which tiers are classified from the perspective of physician. In order to classify the patients’ tier
from the examinees’ point of view, we analyzed the trend of ASD values and established ASD thresholds for
MS subjects’ tier classification. Figures from Fig. 4.10 to Fig. 4.12 show the distributions of frequency counts
over ASD values for each case in male subjects by their age. According to the analysis results, all graphs in
the figure show similar patterns in the distributions of each case except Case 1 and 3. Fig. 4.11 and Fig. 4.12 show almost same shape for the distribution of Case 1 and 3, but Fig. 4.10 depicts different pattern. Because case 1 and 3 are for the subjects having MS, those differences are thought to be caused by the percentage of MS subjects. For middle-aged and old subjects, the incidences of MS are 36.15% and 36.13%, respectively. However, the young subjects show only 16.35% of MS incidence.

![Figure 4.10: Distributions of frequency counts of young-adult male subjects over ASD values.](image1)

![Figure 4.11: Distributions of frequency counts of middle-aged male subjects over ASD values.](image2)
Figure 4.12: Distributions of frequency counts of old-aged male subjects over ASD values.

Figure 4.13 shows the percentages of MS subjects over ASD values for each male subject. As described in Table 4.6, we chose ASD thresholds according to the percentages of MS subjects. Table 4.9 lists the determined ASD thresholds.

Figure 4.13: ASD thresholds for TIER classification of male subjects.
Table 4.9: ASD thresholds for male subjects sub-groups.

<table>
<thead>
<tr>
<th>Patient Tier</th>
<th>TIER(1)</th>
<th>TIER(2)</th>
<th>TIER(3)</th>
<th>TIER(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>ASD&lt;0.80</td>
<td>0.80≤ASD&lt;0.90</td>
<td>0.90≤ASD&lt;0.97</td>
<td>0.97≤ASD</td>
</tr>
<tr>
<td>Middle-aged</td>
<td>ASD&lt;0.69</td>
<td>0.69≤ASD&lt;0.86</td>
<td>0.86≤ASD&lt;0.95</td>
<td>0.95≤ASD</td>
</tr>
<tr>
<td>Old</td>
<td>ASD&lt;0.63</td>
<td>0.63≤ASD&lt;0.81</td>
<td>0.81≤ASD&lt;0.91</td>
<td>0.91≤ASD</td>
</tr>
</tbody>
</table>

Figures from Fig. 4.14 to Fig. 4.16 show the distributions of frequency counts over ASD values for each case in female subjects by their age. As seen in the figures, different distribution patterns of Case 1 and 3 are observed in the all age groups. Unlike analysis results of male subjects, the gaps of frequency counts over ASD values of Case 1 and 3 increase as the age of female subjects increases through Fig. 4.14 ~ Fig. 4.16. The percentage of MS patients in female subjects is significantly increased as the age increase. (5.10%, 25.60% and 48.44% in young, middle-aged, and old female subjects, respectively).

Figure 4.14: Distributions of frequency counts of young-adult female subjects over ASD values.
Figure 4.15: Distributions of frequency counts of middle-aged female subjects over ASD values.

Figure 4.16: Distributions of frequency counts of old-aged female over ASD values.
Figure 4.17 shows the percentages of MS subjects over ASD values for each female subject. Table 4.10 lists the determined ASD thresholds. To sum it up, the rate of MS is higher in female than in male for all ages, especially women are more susceptible to MS as they get older.

![Figure 4.17: ASD thresholds for TIER classification of female patients.](image)

Table 4.10: ASD thresholds for female subjects sub-groups.

<table>
<thead>
<tr>
<th>Patient tier</th>
<th>TIER(1)</th>
<th>TIER(2)</th>
<th>TIER(3)</th>
<th>TIER(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>$ASD&lt;0.81$</td>
<td>$0.81 \leq ASD &lt; 0.89$</td>
<td>$0.89 \leq ASD &lt; 0.97$</td>
<td>$0.97 \leq ASD$</td>
</tr>
<tr>
<td>Middle-aged</td>
<td>$ASD&lt;0.68$</td>
<td>$0.68 \leq ASD &lt; 0.84$</td>
<td>$0.84 \leq ASD &lt; 0.92$</td>
<td>$0.92 \leq ASD$</td>
</tr>
<tr>
<td>Old</td>
<td>$ASD&lt;0.43$</td>
<td>$0.43 \leq ASD &lt; 0.77$</td>
<td>$0.77 \leq ASD &lt; 0.93$</td>
<td>$0.93 \leq ASD$</td>
</tr>
</tbody>
</table>

### 4.4.2 Discussion

The gender differences on the MS related factors have been discussed in many literatures. Among them, Regitz-Zagrosek et al. reviewed on gender differences in the MS. They discovered the gender difference of the components in MS like glucose intolerance pattern, different lipid accumulation pattern in males and
females, and its morphological change for postmenopausal females. These factors affect higher incidence of MS in females group. Also the high development rate of MS in postmenopausal female can be related to sex hormone. The sex hormone which determines the physical and functional characteristics in male and female is thought to be the factor that affects the gender different glucose intolerance and lipid metabolism [61]. Among them, it was revealed that estrogen has an important role in energy homeostasis and metabolic syndrome in both men and women from the studies using estrogen deficient animal models and in estrogen deficient men. A deficiency of estrogen like menopause and/or aging leads to higher incidence of metabolic syndrome in female than in male [62]. The distributions of frequency counts and percentage of MS subjects over ASD value shown in this subchapter indicate difference pattern among gender and age groups, which complies with the literatures. Therefore, we can claim that our proposed method effectively representing the risk of MS disease.

4.4.3 Summary

To effectively manage the temporal changes of MS patients, we proposed a threshold-based patients status classification of MS care based on chronological distance variance. The proposed method classified the status of patients’ tiers having MS using areal similarity degree analysis model and chronological distance variance. We evaluated the proposed method using data obtained from the third Korea National Health and Nutrition Examination Survey among non-institutionalized civilians in the Republic of Korea, which was conducted by the Korean Ministry of Health and Welfare in 2005. The evaluation results showed that our proposed method could classify patients’ temporal changes of MS risk status.
4.5 Summary

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:

- 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 statut, 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 of ASD between consecutive MS risk factors examination results. We evaluated the proposed temporal change model using a sample patient’s examination results over 9 years. 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. We showed the clinical effectiveness of the proposed tier classification using data of a large number of subjects obtained from the KNHANES III. The evaluation results showed that the proposed patients tier classification could help physicians with identifying the major status changes of MS patients’s 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.
Chapter 5. Personalized Healthcare System in Home-Hospital Cloud Environments

This chapter proposes a new healthcare system architecture that integrates the at-home and at-hospital environment to effectively care MS patients. The system supports customizability and dynamic functionality update in a personalized healthcare system by using cloud-based at-home and at-hospital environments integration. The proposed system architecture provides MS risk management functionality based on both MS risk quantification model and temporal progress model. The service broker module within the system supports dynamic provisioning and configuration of personalized at-home healthcare system in cloud environments. We also present the prototype implementation of the personalized healthcare system in home-hospital cloud environments. For easy operation, we have implemented the smartphone user interface and connected the interface with the personalized healthcare system in the testbed. We expect the proposed personalized healthcare system and user interface to be applicable to provide cost effective and personalized chronic disease care.

5.1 Introduction

The increasing number of patients having chronic diseases is becoming an important issue in many countries. Facing the increasing demands and challenges in the area of chronic disease care, various studies on the healthcare system which can, whenever and wherever, extract and process patient data, have been conducted. Chronic diseases are the long-term diseases and require the processes of the real-time monitoring, multidimensional quantitative analysis and the classification of patients’ diagnostic information. Since a patient having chronic disease usually spends most of time with self-management in out-of-hospital environments, particularly at-home environment, and patients have different disease statuses and management requirements, it is important to provide integrated and personalized healthcare services for the patients. A healthcare system for chronic diseases is characterized as an at-hospital and at-home service according to a targeted environment. Both services basically aim to provide patients with accurate diagnoses of disease by monitoring a variety of physical states with a number of monitoring methods, but there are differences between home and hospital environments, and the different characteristics should be considered in order to provide more accurate diagnoses for patients, especially, patients having chronic diseases. However, many studies
aim to develop integrated and computationally efficient healthcare systems, and the development of personalized patients customizable healthcare system is less considered.

To resolve this issue, in this chapter, we investigate a service flow for caring chronic disease patients and firstly develop a new integrated healthcare system architecture that supports both at-home and at-hospital environment. Then, we propose a cloud-based personalized healthcare system for chronic disease care by extending the integrated healthcare system architecture to support patients-centric customizability. The proposed personalized healthcare system provides several chronic disease care services including metabolic syndrome (MS) care services. Finally, the prototype implementation of the proposed personalized healthcare system including the mobile applications for patients is presented. The proposed personalized healthcare system is applicable to providing cost effective MS care at hospital environment and personalized self-management of MS patients at home environment.

5.2 Problem Description

Chronic diseases are increasingly an important concern in u-healthcare systems throughout the world. For example, it is forecasted that clinical expenses for chronic diseases in the U.S. will be 80% of total medical costs and that more than 150 million people may suffer from chronic diseases in 2020 [64]. In response to this challenge, many studies have reported various technical u-health service systems in patient-care monitoring utilizing sensor networks and medical services recently. Particularly, the development of e-health technologies such as mobile computing using dynamic software adaptation techniques [65] or new networking technologies has seen the important elements of chronic conditions based on sensors become a primary issue. In other words, through mobile devices, subcompact sensors, and wireless networks, a health examination is executed and can transmit in real time a patient’s physical data to a medical center. Therefore, these systems enable out-of-hospital health monitoring. Health monitoring of chronic disease patients in out-of-hospital conditions, especially in the home environment, has drawn the attention of healthcare researchers and developers for a long time, because the patients having chronic disease spend most of their time at out-of-hospital environment such as home. Thus, it has become important to provide self-management capability for the patients during daily life [66]. It is known that the problems of conventional healthcare systems have following problems [20].
— Very limitedly available time for problem analysis
— Disconnected environments for patient’s disease management
— Lack of support for user-centric customizable healthcare system
— Computational overhead due to large volume of health information data

Regarding the problems above, there exist many studies to resolve the third problem, i.e. computational overhead, by developing broker systems for healthcare service delivery. Some of the studies can be found in [67-68]. However, the first and second problems have not well investigated so far. Thus, in this chapter, we propose new healthcare system architecture to resolve the first two problems investigated above.

Firstly, from a medical service provider’s point of view, to provide better quality healthcare service for chronic diseases, the following limitations need to be overcome. Patients need a long treatment period with continuous monitoring care. Their condition sometimes may change or worsen unexpectedly. Yet, few existing medical systems provide any alarm about the status of a chronic patient. The usual medical examination processes for disease detection are complicated. Often coping with multiple conditions, chronic disease patients may want to meet multiple physicians at different care sites, but this increases the possibility of risks of errors and poor care coordination. The percentage of reported errors at least doubles among patients seeing four or more physicians compared to those seeing only one or two. Thus, medical systems should generate reliable outcomes for patients with complex chronic conditions [69]. These challenges encourage the improvement of the quality of healthcare systems in terms of hospital demographic factors, clinical technology, information technology, clinical quality, process quality, and hospital financial performance [70-71], the ease of the access to healthcare and healthcare information, and the reduction of the cost of the delivery of healthcare [72-73]. Regarding chronic diseases, Epping-Jordan et al. founded that when patients with chronic conditions receive effective health management within an integrated system, with self-treatment support and regular follow-up, they do better. Thus, healthcare systems need to be prepared to adapt to changing situations, new information, and unforeseen events [74].

Therefore, it is important work to allow chronic patients to manage their own conditions these days, and healthcare systems are required to assist patients’ self-management of their chronic condition by delivering more precise information and suggesting suitable disease management methods [75].
Secondly, from a patient’s point of view, providing a customizable self-management capability for patients having chronic disease is important. Many studies have been done in an effort to provide personalized healthcare services that aim to deliver the ‘right treatment to the right patient at the right dose and at the right time’ [75]. Moreover, the development of an integrated and personalized healthcare system is becoming an important issue in the modern healthcare industry. There have been several studies that have sought to provide integrated healthcare services in home and hospital environments [20]. However, most of these studies have been based on independent healthcare systems that operate in either the home or the hospital; thus, the research has mostly focused on integrating two different healthcare systems seamlessly. Because the requirements of the home and hospital environments are significantly different, cost is a significant issue when building separate healthcare systems and attempting to integrate them seamlessly [64]. However, as the concept of cloud computing has evolved, many studies have reported the potential benefits of cloud computing and have proposed various models or frameworks in an attempt to improve healthcare services [76-77]. These studies have focused on developing a model for ubiquitous healthcare services based on cloud computing, which utilizes a large data archive of clinical data records, decision support systems, and the event-based notification and monitoring system of a typical hospital.

To support customizability and dynamic functionality update in personalized healthcare system, we propose a cloud-based personalized healthcare system, which extends the new integrated healthcare system architecture that supports both at-home and at-hospital environment. The new integrated healthcare system architecture allows creation of patient-specific virtual machines within the hospital cloud environment. Each virtual machine is allocated to a patient and interacts with patient’s healthcare systems such as smartphone, at-home healthcare system gateway, etc. The out-of-hospital self-management results are delivered to the designated virtual machine and the HIMS utilizes the results to investigate the self-management status. The proposed system is providing advanced MS care services in which it combines the MS diagnosis services and temporal progress care services. Using the proposed system, we can provide a patient-centric personalized healthcare service that enables cost effective and personalized chronic disease care.
5.3 Architecture of Personalized Healthcare System in Home-Hospital Cloud Environments

An integrated healthcare system that enables health monitoring and disease management in the home environment has been a major research area for healthcare researchers. Integrated healthcare systems mainly focus on monitoring patients’ health status, detecting and managing potential diseases in the early stage, and managing health problems in daily life [78]. Extending health monitoring from the hospital to the home environment should not be seen as a replication of the same monitoring procedures and methods of the home environment, because the home environment has characteristics that are very different from those of the hospital in terms of medical facility, human resources, the medical knowledge of operator, and other factors. Thus, the approach for simply building the same monitoring architecture as the hospital will dramatically increase the time and human resources necessary for healthcare services delivery. It may also be unacceptable for patients due to its obtrusive and stressing nature. Within the scope of continuity of healthcare, the need to move beyond passive monitoring to efficient mechanisms for personalized healthcare is becoming more and more evident [79]. As was discussed in previous subchapter, one of the emerging requirements for a healthcare system is to provide various health monitoring and disease detection services in the context of each user’s characteristics. The key goal of the healthcare system at a hospital is to detect anomalies in patients’ physiological parameters as accurately as possible and to make diagnoses of potential diseases based on the detected anomalies, whereas the objective of home healthcare is to provide simple and easy methods for monitoring patients’ health status while they carry out the activities of daily life [80].

The functional architecture of the personalized chronic disease care system based on home-hospital cloud environments is depicted in Fig. 5.1. The system developed in this subchapter consists of six key components: a virtual machine manager for personalized home healthcare, the home healthcare system, the home cloud network, the hospital healthcare system, chronic disease care services, and the service broker system.
Figure 5.1: Architecture of the personalized healthcare system for chronic disease in home-hospital cloud environments.
5.3.1 Architecture of Personalized Healthcare System at Hospital Cloud

To overcome the challenges of developing healthcare systems for home environments, it is desirable to support industrial information integration methods such as service-oriented architecture, and multi-tenancy patterns in chronic disease services so that patient-specific healthcare services can easily be integrated into the healthcare system [43][45]. The service-oriented architecture is an appropriate solution for improving integrated healthcare systems, as it is known that the service-orient architecture can provide a unified platform for managing various services, such as data federation, temporal order filtering, and image processing. Also, a service-oriented architecture can offer a reduction in the system complexity, an increase of service extensibility, and good replaceability [17]. Therefore, we designed the architecture of our proposed healthcare system based on the service-oriented architecture (Fig. 5.2). Because the proposed architecture follows the general approach for developing integrated medical information systems in the literature [30][31], we briefly describe the major components of the system and provide a detailed description of chronic disease care services in the following subchapters. The system developed in this research consists of three major components: a personalized user interface, an integrated healthcare server, and chronic disease care services.
Figure 5.2: Architecture of the integrated healthcare system.

A. Personalized User Interface

Our system provides separate user interfaces according to the user types, i.e., the patient and the physician. The patient’s interface displays an abstract and simplified view of the health monitoring results, whereas the physician’s interface shows detailed information, such as patient’s profile and medical history. User
interfaces communicate with the integrated healthcare server by exchanging extensible markup language (XML) based messages. The simple object access protocol (SOAP) was used as the XML-based message binding protocol.

**B. Integrated Healthcare Server**

This component is responsible for the main operation of the system. It consists of six sub-components.

1) **Web Interface Module**: This includes the web application server (WAS), web server, UI repository, and heart disease analysis results module. WAS is used for supporting the web interface of the patient and physician interfaces. If the user interface is a mobile environment, an appropriate application interface for a mobile device, such as a smartphone, is selected from the UI repository and provided.

2) **Service Registry (UDDI)**: Various healthcare services developed by the service-oriented architecture are registered in the Service Registry. The service registry provides naming abilities and translations of names to the web service definition language (WSDL). A more detailed description of the operational procedures of the chronic disease care services is presented in the following subchapters. Among the registered healthcare services, appropriate services for patients’ characteristics, such as the chronic disease and the service environment, are executed to provide personalized healthcare service.

3) **Communication and Monitoring Module**: This module takes care of communication among the healthcare service participants, such as patients, physicians, and informants, and collects various types of measured data of patients on either a real-time or a non-real-time basis. The collected data are then transferred to the decision support tool.

4) **Decision Support Tool**: This identifies patients’ chronic diseases and their status using MS risk model proposed in Chapter 3 and 4.

5) **Patient-specific Disease Care Service Module**: This component receives information about patients, including the patient profile and chronic disease, and analyzes patients’ measured health data. In order to perform the analysis, appropriate healthcare services registered in the Service Registry are
utilized. The analyzed results are moved to the Web Interface Module and are finally delivered to the user interfaces of the patients and physicians.

6) Integrated Healthcare System Middleware: Each type of module in the server can communicate with healthcare databases using middleware.

C. Virtual Machine as a Hospital Cloud Environment for Personalized Home Healthcare

This component is responsible for the main operation of the system. It creates a virtual machine in the hospital cloud environment and customizes the virtual machine according to the patient’s characteristics. This component receives information about the patient, including the patient profile and the status of the chronic disease. It then uploads the disease care services to the virtual machine based on the patient’s service class and disease status and configures the service parameters according to the patient’s health status. The disease care services are installed to the virtual machine by downloading services from application and service repository. After installing the chronic disease care services, the information about the patient’s mobile device for use with the home healthcare service and the synchronization method for health monitoring data is configured. The creation and configuration of the virtual machine is done via the integrated healthcare system at hospital cloud.

D. Decision Support Tool and Visualization Module

The decision support tool is based on the patient status classification method to identify and classify the patients’ disease statuses [20]. In addition to the conventional patient status classification method, the functionality of decision support tool is extended to support risk prediction and temporal progress analysis of chronic disease, especially MS [56][63]. An additional MS care module is further added to the decision support tool in order to support chronic disease status visualization functionality on mobile devices using a Java-based widget interface. Through the mobile widget on a device, we can use data in conjunction with the medical diagnosis system as a progressive medical service. The prototype implementation of the disease status visualization interface will be presented in the following Subchapter 5.4.
E. Chronic Disease Care Services

Currently our system provides two types of chronic disease care services: the MS Risk Quantification Service and the MS Temporal Progress Care Service. Also, it is planned to extend the system to support additional services such as the Mental Stress Factor Identification Service, the e-Re Number-Estimation-based Heart Disease Detection Service, and the ECG Signal-Analysis-based Arrhythmia Detection Service. The references of the services are registered in the service registry. Our proposed integrated healthcare system architecture supports two types of chronic disease care services. Also, it is planned to adopt four additional chronic disease care services in the future. Each chronic disease care service can be developed in a separate module and it is registered to Service Registry module using UDDI. To improve the compatibility between the chronic disease care services and our proposed integrated healthcare system, XML-based messaging and data exchange are utilized.

5.3.2 Patient-Centric Chronic Disease Care Services Provided by Personalized Healthcare System at Hospital Cloud

As we described in the previous subchapter, our proposed integrated healthcare system architecture supports four types of chronic disease care services. Each chronic disease care service can be developed in a separate module and it is registered to Service Registry module using UDDI. To improve the compatibility between the chronic disease care services and our proposed integrated healthcare system, XML-based messaging and data exchange are utilized. This subchapter describes the overview and Petri-net based service flow of each chronic disease service.

A. Chronic Disease Care Service: MS Risk Prediction Service

MS has emerged as an important issue of public healthcare in many countries. MS refers to a clustering of CVD related risk factors whose underlying pathology is related to insulin resistance. The risk factors include insulin resistance, obesity, dyslipidemia, and hypertension, and it is known that they increase the risk for CVD and T2DM [1-2][67]. Therefore, the importance of MS is that it helps to identify individuals at high risk for both CVD and T2DM. There has been much effort to establish diagnostic criteria for MS, but the cur-
rent diagnostic criteria of MS have weaknesses, such as binary decision based on diagnostic criteria, equal weight among risk factors, and difficulty in estimating the temporal progress of the risk factors [1-2][4]. To resolve these problems, Jeong et al. developed a risk quantification model for MS, which was based on areal similarity degree analysis between weighted radar charts comprising metabolic syndrome diagnostic criteria and examination results of risk factors [56]. The MS risk quantification and prediction model is integrated into our proposed personalized healthcare system to predict the risk of having MS in the future. Fig. 5.3 shows the operational procedure of the MS risk prediction service in the proposed system. The risk prediction service is provided as one of the chronic disease care services in our proposed cloud-based personalized healthcare system.

![Operational procedure of the MS risk prediction service](image)

**Figure 5.3:** Operational procedure of the MS risk prediction service.

**B. Chronic Disease Care Service: Temporal Progress Identification Service**

A patient’s medical examination results and profile information are used to calculate the quantified risk of having MS. Then, distance function based on chronological clustering method is calculated. If the distance result exceeds the designated control ranges of the patient, the patient is likely to have major changes at disease status. Therefore, the medical examination results and the patient’s profile information are delivered
to the physician for further investigation and the patient is instructed to consult the physician. Fig. 5.4 shows the operational procedure of the temporal progress care service of MS, which is based on the chronological distance analysis method proposed in [63].

Figure 5.4: Operational procedure of the temporal progress care service for the MS.

C. Chronic Disease Care Service: Sensitivity-Level Based Patient Tier Classification Service

To provide patient-centric personalized chronic disease care services, it is necessary to support different service levels according to the patient’s disease status. Thus, different types of services are provided to patients based on the patient tier, as shown in Table 5.1. The basic service level is mapped with TIER(1) and TIER(2) patients that have low risk for MS. The TIER(3) patient group is assigned to the standard service level and TIER(4) group for the premium service level. To determine patient tier effectively, we extend internal module of decision support tool to support service level-based as shown in Fig 5.5. The examination
results from the disease and complications identifier module can be generated at each individual hospital. The results obtained from the remote hospitals can be integrated with web-based system and can be calculated with the mapping module in the health risk quantification module support the temporal progress of MS identification method presented in this subchapter. By using the calculation results of the health risk quantification module, the patient status classifier determines patients’ chronic disease status.

Figure 5.5: Health risk quantification process.

Table 5.1: Three service levels for chronic disease patients.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Service Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Basic) Diagnosis Service</td>
<td>Basic TIER(1)&amp; TIER(2)</td>
</tr>
<tr>
<td></td>
<td>- Patient Profile (S1)</td>
</tr>
<tr>
<td></td>
<td>- Disease Profile (S2)</td>
</tr>
<tr>
<td></td>
<td>- Metabolic Syndrome Risk Prediction (S3)</td>
</tr>
<tr>
<td></td>
<td>- Chronic Disease Temporal Progress Identification (S4)</td>
</tr>
<tr>
<td></td>
<td>- Physician Comments (S5)</td>
</tr>
<tr>
<td>Advanced Diagnosis Service</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Result Report</td>
<td>- Web-Portal (S14)</td>
</tr>
<tr>
<td></td>
<td>- E-Mail (S15)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiences</td>
<td>- Physician With Less than 5 years of experience</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td>6 months</td>
</tr>
<tr>
<td>Monitoring Interval</td>
<td>- Every month</td>
</tr>
<tr>
<td>Quality</td>
<td>- Within 1 week</td>
</tr>
<tr>
<td></td>
<td>- Stable Resources</td>
</tr>
<tr>
<td></td>
<td>- Low Reliability</td>
</tr>
<tr>
<td>Cost</td>
<td>1 Unit Cost</td>
</tr>
</tbody>
</table>
D. Chronic Disease Care Service: Informant-driven Mental Stress Factor Identification Service

The mental stress level and heart disease, such as tachycardia or bradycardia, demand close patient observation and constant patient care. In order to manage such patients, a physician needs to know the heart rate and stress response inventory (SRI) scores. Then, the physician may investigate the physiological relationships between heart rate and each SRI score to identify mental stress factors that may induce heart rate changes. Several studies have been performed to identify the relationship between the mental stress factor and heart rate changes [86-88]. We have adopted the stress factor identification method presented in [87-88] to the at-home-hospital environment so that physicians can investigate the relationship between mental stress factors and chronic disease, as shown in Fig. 5.6. The stress factor identification method is based on multiple linear regression models between heart rates and each SRI score.

![Diagram of the Chronic Disease Care Service](image)

Figure 5.6: Informant-driven mental stress factor identification service.

22 SRI questions ($Q_{SRI}$) are transmitted periodically from the informant to the patient’s mobile device. The patient replies to the SRI questions and $n$ ECG signals are transmitted from the patient through the mobile device to the informant at the same time. The transmitted SRI responses and ECG signals are saved in the informant’s database. Then, if the patient’s heart rate changes, the informant notifies the stress heart information system and the physician to identify the mental stress factors ($Q_{MLRM}$) with multi-linear regression (Fig. 5.7) [88].
E. Chronic Disease Care Service: e-Re Number Estimation based Heart Disease Detection Service

Cardiac arrhythmia is life-threatening medical emergency that can result in cardiac arrest and sudden death. According to a medical report by the American Heart Association 2010, approximately 55% of heart disease patients die due to arrhythmia. Serious cases of arrhythmia, such as ventricular tachycardia or fibrillation, are mostly induced by vortex-like reentrant electric waves in cardiac tissue. In this study, we have integrated a new service for arrhythmia management based on a biomarker index of cardiac arrhythmia in our integrated healthcare system. The index, termed ‘e-Re’, was introduced in our previous paper. It is partially based on the Reynolds number of the fluid dynamics [89]. The index and Reynolds number are similar in form, physical meaning, and function. It was shown that electrical wave patterns in cardiac tissue can be classified according to the magnitude of the index, just as fluid flow patterns are determined by Reynolds numbers in fluid dynamics. The e-Re number can categorize reentrant arrhythmia in the human heart. Specifically, if the cardiac electrophysiological condition of the e-Re number exceeds a critical threshold value, reentrant tachycardia or fibrillation is easily generated by some abnormal ectopic beat.
The integrated healthcare system can use the $e$-$Re$ number to monitor the arrhythmic risk of patients with real-time ECG signals. To calculate the $e$-$Re$ number, some parameters of the patient, such as the characteristic frequency and length of the patient’s heart and electric conduction coefficient of the heart, are required. The characteristic length of the heart is determined by the muscle volume of the left and right ventricles. The muscle volume of the heart can be measured by ultrasound or with a computed tomography device once a year or once every half year, as it does not change easily within short term shorter than one year. The conduction coefficient of the cardiac tissue is estimated from duration of the P wave or the QRS complex, which can easily be measured using an electrocardiography device. The characteristic frequency indicates the heart rate. Therefore, the integrated healthcare system can provide more detailed and specific information about an arrhythmic patient by obtaining the $e$-$Re$ number in addition to the real-time electrocardiography signal. Fig. 5.8 shows the service flow for the $e$-$Re$ number-estimation-based heart disease detection service.

Figure 5.8: Operational procedure for the $e$-$Re$ number estimation-based heart disease detection service.

**F. Chronic Disease Care Service: Microscopic Perspective Mitochondria-level Metabolic Disorder Prediction Service**

Metabolic disorder occurs when abnormal chemical reactions in human body disrupt this process. It was well known that the mitochondria are “the body engine” for energy production, enabling life activity. Their dysfunction can induce metabolic disorder [90]. According to epidemiologic and clinical observations
[91-92], a decrease in the mitochondrial deoxyribonucleic acid (mtDNA) copy number is responsible for type II diabetes mellitus. Thus, it is critical to measure the capability of mitochondria energy production for an early diagnosis of metabolic disorder or T2DM.

Recently, we developed a prediction system for metabolic syndrome by adding the information of mitochondria metabolism as measured using nano-techniques [90]. In the study [90], it was shown that mitochondria metabolism could be evaluated by the capability of energy production, which is closely related to the cellular temperature. We also discussed a sensor-integrated system model for metabolic syndrome predictions with a workflow system. This model measures not only the cell temperature variation using an invasive method, but also involves a controlling simulation for metabolic syndrome predictions.

To predict metabolic disorder, the authors designed a hybrid metabolic syndrome analysis system (Fig. 5.9) that has high precision sensor units, a network interface to deliver the analysis results of human cells, annotation data from a public hospital, and metabolic data.

Figure 5.9: Operational procedure for the mitochondria-level metabolic disorder prediction service.
5.3.3 Architecture of Patient-centric Personalized Home Healthcare System at Home Cloud

A. Home Healthcare System

The home healthcare system is configured by migrating the virtual machine created in the hospital cloud environment to the patient’s home cloud. The home healthcare system communicates with the virtual machine at the hospital via the Sync-Server in the integrated healthcare system at the hospital. When a reconfiguration of the patient’s home healthcare system is needed, for example a change in the patient’s service level or an update of the disease care services, the virtual machine at the hospital will be reconfigured and the updated virtual machine will be migrated to the patient’s home healthcare system. Thus, it is possible to provide personalized healthcare services to patients.

When the home healthcare system receives health monitoring data from the patient’s mobile device, it analyzes the data using the disease care services installed during the configuration of the home healthcare system by the hospital. The home healthcare system then delivers the analysis results to the patient, the caregiver, and to the family members’ mobile devices. For a premium service class patient, the home healthcare system can deliver the analysis results to mobile devices at a remote site. At the same time, the home healthcare system transfers the sensed data to the virtual machine created in the hospital’s cloud environment, and it can synchronize the patient’s disease status and monitoring data. The home healthcare system communicates with the virtual machine at the hospital via broadband and the data communication utilizes international standards such as the ANSI health level 7 reference information model (HL7 RIM), the clinical document architecture (CDA) v2, and the extensible markup language (XML) to improve the interoperability of the system and to ensure integration with existing healthcare information systems.

B. Home Cloud Service Environment

At the patient’s home cloud service environment, the patient’s mobile device collects health monitoring data from the portable sensors. Sensor data can be collected via an IEEE 802.15.4 wireless connection, which was selected due to its wide deployment in biomedical devices. The mobile device processes the re-
ceived data to detect any mis-positioning of the sensors and then forwards the data to the home healthcare system through a Wi-Fi connection when the patient is within the home network, or via the mobile network when the patient is out of home the network to be further analyzed using the chronic disease services. The availability of multiple communication paths ensures the good adaptability of the system in a multitude of operating areas and improves its fault tolerance.

C. Service Broker for Personalized Healthcare System

The service broker is responsible for allocating resources in cloud to the patient’s home cloud virtual machine. Since our proposed system supports different service level among patients, the service broker discovers resources and allocates appropriate resources to tasks that are divided into workflows. The service broker architecture is based on an active workflow control scheme, which includes a policy manager, workflow manager, and resource manager to guarantee service level agreements (SLAs) of chronic disease care services. Fig. 5.10 shows the functional architecture of the service broker.

Figure 5.10: Service broker module for personalized healthcare system in cloud environment.

SLAs for chronic disease care services are delivered to the broker through the integrated healthcare system interfaces. Then, the requests from the chronic disease care services are submitted, and parsed into
tasks and dependencies in the workflow management engine. For each application submitted, the workflow manager analyzes the input application and workflow. Then the manager divides the application into tasks corresponding to workflow. The workflow scheduler maps the tasks into the workflow using policy decision to achieve optimized mapping based on the SLA requirements. After receiving the mapping information from the workflow manager, the resource manager schedules tasks and discovers available resources which can guarantee quality of service. The resource selection manager chooses resources for executing schedules tasks and allocates tasks to the selected resources. At the same time, the resource manager creates task and resource managers to manage and monitor the status of tasks and resources. The resource manager is responsible for monitoring the resource status, generating the list of available resources, and dispatching the tasks into the resources in hospital cloud environment.

5.4 Evaluation and Prototype Implementation

This subchapter presents evaluation results of the proposed personalized healthcare system architecture using Markov-process-based cost analysis. Then, the prototype implementation results of the healthcare system are also presented.

5.4.1 Prototype Implementation

This subchapter presents the prototype implementation results of the proposed personalized healthcare system for cloud environments. We built a prototype system on the experimental testbed as shown in Fig. 5.11. Firstly, a patient downloads at-home healthcare system application from service repository to the patient’s mobile devices such as smartphone. Then, the patient accesses the virtual device that is synchronized with a virtual machine within the hospital cloud and allocated to the patient. The service broker is responsible for creating and configuring the virtual machine within the hospital cloud. The virtual machine is corresponding to the patient’s home healthcare machine, which is a virtual device in at-home cloud environment. The broker installs the chronic disease care services customized to the virtual machine. The chronic disease care services are customized to the patient’s status by means of service level agreement between the patient and the service broker. Therefore, when an administrator in the hospital needs to update the patient’s home healthcare system, the administrator makes an update on the virtual machine within the hospital cloud. Then, the virtual
machine within the hospital will be synchronized with the virtual device in the patient’s home cloud environment. By doing so, the hospital can effectively manage the status of out-of-hospital patients.

Fig. 5.11: Mobile application for personalized chronic disease care service.

Fig. 5.12 shows the prototype implementation of a patient’s home healthcare devices. Our prototype system is implemented on Android 4.2 platform and supports two chronic disease care services presented in this paper, metabolic syndrome risk prediction service and temporal progress care service, respectively [56][63]. We will extend our system to support other chronic disease services such as ECG analysis services presented in [20].
A. Clinical Applicability for Temporal Disease Status Care

The implemented application on mobile device helps with managing the temporal changes in MS risk. For example, as shown in the Fig. 5.13, we can consider the scenario that a young male individual regularly examines the values of MS risk factors. The mobile application supports the calculation of the ASD, which quantifies the risk of having MS in the future [56]. If the ASD value of the person exceeds the threshold of young male at some time, the status of each risk factor will be investigated in detail and risk factors approaching the thresholds will be managed by physician. Also, during the following regular examinations, not only calculation of ASD, but also generation of the visualization chart of a patient’s health status shown in Fig. 3.4 will be used to analyze the temporal changes of the person. Once the ASD value of the subject be-
comes below the threshold, only the ASD value will be used to regularly manage the person [56].

Furthermore, since patients having chronic diseases such as MS typically spend most of time out of hospital environment between regular health examinations, it will be helpful for patients to perform self-management with simple applications for personal portable devices such as smartphone. By regularly analyzing the ASD value and the visualization chart of MS risk factors, it is possible to manage the temporal progress of MS [63].

Figure 5.13: Temporal changes of ASD value during regular health examinations.

5.5 Summary

Chronic diseases are the long-term diseases and require the processes of the real-time monitoring, multidimensional quantitative analysis and the classification of patients’ diagnostic information. A healthcare system for chronic diseases is characterized as an at-hospital and at-home service according to a targeted environment. Both services basically aim to provide patients with accurate diagnoses of disease by monitoring a variety of physical states with a number of monitoring methods, but there are differences between home and hospital environments, and the different characteristics should be considered in order to provide more accurate diagnoses for patients, especially, patients having chronic diseases. Further, as the number of patients having chronic diseases increases, developing a personalized healthcare system is becoming a major issue in many countries. Key requirements for personalized healthcare system include consideration for separate environments, reduction of computational overhead for health information data processing, and support for
customizable healthcare system. However, many studies aim to resolve the first two requirements, and patient customizability is less considered.

In this chapter, we proposed a new healthcare system architecture that integrates the at-home and at-hospital environment. Further, to support customizability and dynamic functionality update in a personalized healthcare system, we extend the integrated healthcare system architecture and proposed a cloud-based personalized healthcare system for chronic disease care, particularly MS. The proposed personalized healthcare system provides several chronic disease care services including MS risk quantification and temporal progress care services. The service broker module within the system supports dynamic provisioning and configuration of personalized at-home healthcare system in cloud environments. We also presented the prototype implementation of the personalized healthcare system in home-hospital cloud environments. We expect the proposed personalized healthcare system to be applicable to provide cost effective and personalized chronic disease care.
Chapter 6. Conclusion

Facing the increasing demands and challenges in the area of chronic disease care, a number of studies on the healthcare system which can, whenever and wherever, extract and process patient data, have been conducted. Chronic diseases are the long-term diseases and require periodic monitoring, multidimensional quantitative analysis, and the classification of patients’ diagnostic information. Among the chronic diseases, MS that refers to a clustering of specific cardiovascular disease risk factors whose underlying pathology is thought to be related to insulin resistance is one of the major chronic diseases in many countries including Korea, because of its relationship with the incidence of CVD and T2DM. Different from acute disease, chronic disease such MS requires long term care and its temporal information plays an important role to manage the status of disease. Health monitoring in out-of-hospital conditions, especially in the home environment, has drawn the attention of healthcare researchers and developers for a long time, because patients having chronic disease such as MS typically spend most time at home environment.

In response to those increasing requirements, there have been a number of previous studies regarding MS, but most of them were focused on investigating the relationship between MS risk factors and the incidence of other chronic diseases such as CVD, CHD, and T2DM. Up to our knowledge, there has been no previous literature about quantifying the risk of MS, which is essential for predicting the incidence of MS in the future. To achieve this objective, it is imperative to overcome the well-known limitations of MS diagnostic definitions. There has been much effort to establish diagnostic criteria for MS, but it is known that current diagnostic criteria of MS have following weaknesses:

— 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

This dissertation proposed the novel MS risk quantification model, which resolved the weaknesses of MS diagnosis methods and showed the validity of the model using extensive clinical evaluation using a large number of sample health examination data. In this dissertation, we first proposed a risk quantification model for MS, which is based on areal similarity degree analysis between weighted radar charts consisting of MS diagnostic criteria and examination results of MS risk factors of a subject. The clinical effectiveness of the
proposed model was evaluated using data of a large number of subjects obtained from the KNHANES III. The evaluation results showed that the proposed model can quantify the risk of MS and effectively identify a group of subjects who can be classified into a potential risk group for having MS in the future. Using the proposed model, we could early identify potential MS patients and monitor the temporal change of the patients’ statuses. Then, we proposed a patient-specific chronic disease care system using a chronological clustering method to analyze the temporal progress of chronic disease. We also presented a patient tier classification method based on the sensitivity level for accepting or controlling the changes of patient’s disease status. The proposed system provides personalized chronic disease care services according to the classified patient tiers. Through these new technologies, we can design new application services, such as analyzing long-term trends of patient’s disease status, a knowledge-based decision support tool for cardiovascular disease, and a system which predicts mitochondria-level metabolic disorder. The clinical effectiveness of the proposed model was evaluated using data of a large number of subjects obtained from the KNHANES III. The evaluation results showed that the proposed model can be used to analyze the temporal progress of chronic disease, especially, the risk of MS. By using the temporal model, we could effectively manage patients with MS or patients at risk of developing it. Finally, we proposed a new healthcare system architecture that integrates the at-home and at-hospital environment to effectively care MS patients. The system supported customizability and dynamic functionality update in a personalized healthcare system by using cloud-based at-home and at-hospital environments integration. The proposed system architecture provides MS risk management functionality based on both MS risk quantification model and temporal progress model. The service broker module within the system supported dynamic provisioning and configuration of personalized at-home healthcare system in cloud environments. We also presented the prototype implementation of the personalized healthcare system in home-hospital cloud environments.

The primary contribution and applicability to healthcare and biomedical engineering field is that it proposed an innovative method to identify potential patients with having high risk of MS in advance so that physicians and patients could proactively manage health status and reduce time and medical expenditure for caring the MS. To achieve the objective, we developed a novel MS risk quantification model based on ASD analysis, a temporal progress model based on chronological clustering methodology, and an integrated healthcare architecture for home-hospital integrated healthcare system for MS care.
References


[28] National Cholesterol Education Program - Third Adult Treatment Panel (NCEP ATP III)


[71] L. Li and D. Collier, “The role of technology and quality on hospital financial performance: An explor-


[83] A. Shaikh, M. Memon, N. Memon, and M. Misbahuddin, “The Role of Service Oriented Architecture in Telemedicine Healthcare System,” in *Proc. 2009 Int. Conf. on Complex, Intelligent and Software In-


[125] Health Insurance Review & Assessment Service (HIRA), Available at http://www.hira.or.kr/main.do

[126] Seoul National University Hospital, Available at http://www.snuh.org/pub/snuh/sub01/sub05/


요 약 문

클라우드 통합형 헬스케어 환경에서 대사중후군 관리를 위한 의료의사결정시스템에 관한 연구

만성질환 관리 분야의 지속적인 요구 및 도전에 직면하여, 환자의 의료 데이터를 언제, 어디서나 활용하고 처리할 수 있는 헬스케어 시스템에 대한 많은 연구가 수행되었다. 만성질환은 장기적인 관리가 필요한 질병으로 환자의 진단 정보의 분류, 상태의 추적적인 검사, 다양한 측면에서의 결과 분석 등이 필요하다. 이러한 만성질환 중 대사중후군은 심혈관계 질환의 위험요소 중 병리적 특성이 인슐린 저항성과 관계가 있는 요소들의 군집현상으로 지칭되는 질환으로, 우리나라를 비롯한 많은 나라에서 주요한 만성질환 중 하나로 인식되고 있다. 대사중후군은 심혈관계 질환이나 재발당뇨병의 발현과 관련이 있다는 점에서 중요성이 부각되고 있다. 급성질환과 달리 대사중후군과 같은 만성질환은 장기적이며 시간에 따른 변화도 관리가 질환의 효과적인 관리에 중요할 요소이다. 환자들 특히 만성질환 환자들의 가정과 같은 병원 이외의 환경에서의 대부분의 시간을 보내며 자가관리를 수행하기 때문에, 이러한 환경하에서의 효과적인 건강 상태 모니터링은 헬스케어 관련 분야의 연구 개발자의 주요한 연구 분야 중 하나로 대두되어 왔다. 이러한 요구에 부응하여, 대사중후군 분야의 많은 연구가 진행되어 왔으나, 기존의 연구들은 대부분 대사중후군의 위험요소와 심혈관계 질환, 당뇨병과 같은 다른 만성질환의 발병간의 관계를 규명하는 관점으로 수행되어 왔다. 그러나, 미래의 대사중후군은 병발 가능성을 예측하기 위해서 필수적인 대사중후군의 위험도를 계량화하고 평가하는 방법에 대한 진행연구는 현재까지 보고되지 않고 있다. 이러한 문제를 해결하기 위해서는 기존의 대사중후군 진단방법의 문제점들의 해결이 필수적이다. 그 동안 대사중후군의 진단기준을 수립하기 위하여 많은 연구와 논의가 진행되어 왔으나, 현 대사중후군의 진단기준은 진단기준 상의 위험이나 이질한 상이한 중요도가 고려되지 않고, 진단기준 값을 넘었는지 넘지 않았는지를 기반으로 질환의 유무를 파악함에 따라 대사중후군 비진단자의 대사중후군 발생 위험도 추정이 어려우며, 대사중후군 상태의 시간의 경과에 따른 상태 변화의 관리가 어려운 것과 같은 대표적인 문제점을 가지고 있는 것으로 널리 알려져 있다. 본 논문에서는 기존 대사중후군 진단방법의 문제를 해결하기 위해서 대사중후군 위험도 계량화 모델을 제안하고, 대규모 그룹의 진단검진 결과를 활용하여 계산 모델의 임상적 유용성을 집중적으로 평가하였다.

본 논문에서는 우선 대사중후군 위험요소별 진단기준 값과 검진자의 진단결과를
이용하여 각각 구성한 가중치 기반의 레이터 차트상에서 두 레이터 차트의 면적 유사도 분석을 활용한 대사증후군 위험도 계량화 모델을 최초로 제안한다. 제안된 위험도 모델의 임상적 유용성은 보건복지부 산하의 질병관리본부에서 제공하는 국민건강영양조사 3기의 원자료를 활용하여 평가하였다. 분석 결과, 제안된 모델은 대사증후군 위험도를 계량화 할 수 있으며, 장래에 대사증후군 발병의 가능성이 있는 잠재적 위험군을 판별하는데 활용될 수 있음을 확인하였다. 제안된 모델을 통해 잠재적 대사증후군 환자를 조기에 판별하고 시간의 경과에 따른 환자의 상태를 모니터링 하는 목적으로 활용될 수 있다.

그 후, 연대순의 클러스터링 기법을 활용한 만성질환의 시간에 따른 시변 상태 변화도를 분석할 수 있는 모델을 제안하고, 제안된 모델을 탑재한 환자 맞춤형 만성질환 관리 시스템을 개발하였다. 또한, 환자의 건강 상태의 변화량을 관찰하고 동체하기 위한 인공지능 기반의 환자 상태 분류 방법도 제안하였다. 개발된 만성질환 관리 시스템은 환자의 상태 분류에 따라 맞춤형 만성질환 관리 서비스를 제공하게 된다. 제안된 만성질환 시변 상태 변화도 분석 모델 및 환자 상태 분류 방법 또한 국민건강영양조사 3기의 원자료를 이용하여 임상적 의의를 평가하였다. 평가결과 제안된 모델이 대사증후군과 같은 만성질환의 시간에 따른 변화도를 효과적으로 분석할 수 있음을 확인하였다. 제안된 시변 상태 변화도 모델을 활용하여 대사증후군 환자 또는 잠재적 위험군에 대해서 건강 상태를 효과적으로 관리할 수 있다.

마지막으로, 본 논문에서는 대사증후군 환자의 효과적인 관리 위한 가정-병원 통합 환경을 지원하는 새로운 헬스케어 시스템 구조를 제안하였다. 제안된 시스템은 클라우드 기반의 가정-병원 통합 기법을 활용하여 환자 맞춤형 기능 제공을 지원하는 헬스케어 시스템이다. 제안된 시스템 구조는 대사증후군 위험도 모델 및 시변 상태 변화도 모델에 기반하여 대사증후군 위험도 관리 서비스를 제공할 수 있다. 시스템 내의 서비스 모듈은 클라우드 환경에서 홍 클라우드에 구성된 가정용 헬스케어 시스템의 동적 설정 및 서비스 제공을 지원할 수 있다. 마지막으로 본 논문에서 제안한 가정-병원 통합형 환자 맞춤형 헬스케어 시스템의 프로토타입 구현 결과를 제시하였다.

본 논문의 주요 기여 및 생명공학 분야의 활용도는 의사와 대사증후군 고위험군 환자들의 건강 상태를 전자적으로 관리하고 대사증후군의 관리를 위해 지출되는 시간과 의료비 지출을 절감할 수 있는 혁신적인 대사증후군 위험도 모델을 제안하고 임상적 유용성을 평가하였다는 것에 있다. 이러한 목적을 위해 본 논문에서는 혁신적인 대사증후군 위험도 계량화 모델과 시변 상태의 변화도를 분석할 수 있는 모델을 개발하고 대사증후군 환자의 효과적인 관리를 위한 가정-병원 통합형 헬스케어 시스템의 구조를 개발하여 상기 모델들을 지원하는 의료의사결정시스템을 탑재함으로써 의사의 환자 진단 및 시 활용 가능하도록 하였다. 본 논문의 구성은 다음과 같다. 2장에서는 만성질환 관리 사이클의 특성 및 주요 만성질환의
유병률을 분석한다. 분석된 질환 중 본 논문에서 다루는 대사증후군의 개념 및 진단기준에 대해 상세히 검토하고 효과적인 만성질환 관리를 위한 헬스케어 시스템 기술 분야의 최신 연구동향을 살펴본다. 분석 결과를 기반으로 현재의 대사증후군 관리상의 주요 문제점을 도출한다. 그 후, 효과적인 대사증후군 환자의 관리를 위한 가정-병원 통합 환경의 헬스케어 서비스 시나리오를 제시한다. 3 장에서는 현재의 대사증후군의 문제를 극복하기 위하여 변적 유사도 분석 기법을 기반하여 잠재적 대사증후군 환자를 판별할 수 있는 대사증후군 위험도 계량화 모델을 제안하고, 제안 모델의 임상적 유용성을 5355 명의 건강검진 결과를 이용하여 평가하였다. 4 장에서는 대사증후군 환자의 시간에 따른 질병 상태의 변화를 관리하고 통제하기 위한 시변 상태 변화도 모델을 제안하고, 임상적 의의를 평가하였다. 5 장에서는 만성질환의 관리적 특성에 기반한 클라우드 기반의 가정-병원 통합형 헬스케어 시스템 구조를 제안하고 프로토타입 구현 결과를 제시하였다. 마지막으로 6 장에서 본 논문의 결론 및 제한사항과 향후 연구 주제에 대해 기술한다.

핵심어: 클라우드, 헬스케어, 만성질환, 대사증후군, 의료의사결정시스템, 의료정보
Academic Activities

1. Book


2. International Journal Papers


3. International Conference Papers


[2] Sangjin Jeong, Seonghwan Kim, Daesun Kim, Chan-Hyun Youn, Yong-Woon Kim, “A Personalized Healthcare System for Chronic Disease Care in Home-Hospital Cloud Environments,” in Proc. IEEE In-


4. Domestic Journal Papers


5. International Patents

6. Domestic Patents


7. International Standards


8. Other Activities

[1] Member of Technical Program Committee on The Sixth IEEE International Conference on Ubiquitous and Future Networks (ICUFN 2014), 2014.

[2] Member of Technical Program Committee on IEEE International Conference on Communications (ICC) 13 WS - 2nd Workshop on Clouds, Networks and Data Centers - A holistic approach towards an integrated service provider infrastructure, 2013.


