

2008

Design and implementation of reverse risk assessment software

Igor S. Tilinin
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Tilinin, Igor S., "Design and implementation of reverse risk assessment software" (2008). *Master's Theses*. 3524.
DOI: <https://doi.org/10.31979/etd.4826-kzrs>
https://scholarworks.sjsu.edu/etd_theses/3524

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

DESIGN AND IMPLEMENTATION OF REVERSE RISK ASSESSMENT
SOFTWARE

A Thesis

Presented to

The Faculty of the Department of Computer Engineering

San Jose State University

In Partial Fulfillment

of the Requirements of the Degree

Master of Science

by

Igor S. Tilinin

May 2008

UMI Number: 1458158

Copyright 2008 by
Tilinin, Igor S.

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 1458158

Copyright 2008 by ProQuest LLC.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

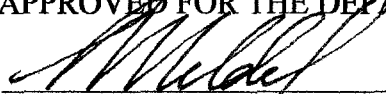
ProQuest LLC
789 E. Eisenhower Parkway
PO Box 1346
Ann Arbor, MI 48106-1346

© 2008

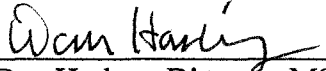
Igor S Tilinin

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF COMPUTER ENGINEERING



Dr. Sigurd Meldal, Chair, Department of Computer Engineering



Dan Harkey, Director, MS in Software Engineering



Russel Smith, Faculty Member, Department of Computer Engineering

APPROVED FOR THE UNIVERSITY

 06/18/08

ABSTRACT

DESIGN AND IMPLEMENTATION OF REVERSE RISK ASSESSMENT SOFTWARE

by Igor S. Tilinin

This thesis addresses the problem of reverse risk assessment of diet-related chronic diseases. The proposed solution is based on a mathematical model which relates diet risk factors with the probability of contracting a chronic disease and a software system capable of generating a large number of possible food intake scenarios (patterns). The developed reverse risk assessment system features three-tier architecture and is built using open source technologies and components like JavaServer Faces, Apache Tomcat and MySQL relational database. A part of the system is a web-based application which allows the user to search and analyze food intake patterns pertaining to different diet types including vegan, vegetarian, Mediterranean and traditional American. Analysis of pattern sets generated by randomly selecting food items from the USDA food composition database reveals that the traditional American diet is characterized by the highest risk for contracting various cardiovascular diseases at a later age.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Sigurd Meldal for giving the opportunity to perform this research work at the Department of Computer Engineering, Dan Harkey for valuable discussions, support and constant attention to the progress of this project. Special thanks are given to Russell Smith who was the first to draw the author's attention to the risk management questions.

Table of Contents

1.	Reverse Risk Assessment and Diet-Related Chronic Diseases	1
1.1	Introduction.....	1
1.2	Direct versus Reverse Risk Assessment	3
1.3	Problem of Reverse Risk Assessment and Diet-Related Chronic Diseases.....	8
1.4	Disease Risk Factors	11
1.5	Diet-Related Risk Factors	14
1.5.1	Fat	15
1.5.2	Cholesterol	17
1.5.3	Fiber	19
1.5.4	Vitamins.....	19
1.5.5	Minerals	24
1.5.6	Sugar	27
1.5.7	Salt	28
1.5.8	Alcohol.....	30
1.6	Diet-Related Chronic Diseases	31
1.6.1	Coronary Heart Disease	32
1.6.2	Stroke	36
1.6.3	Hypertension	38
1.6.4	Obesity	40
1.6.5	Diabetes Mellitus	41

1.6.6	Osteoporosis.....	43
1.6.7	Dental Caries.....	44
1.7	Overview of Existing Food Pattern Calculators	47
2.	System Design	51
2.1	System Requirements Overview.....	51
2.2	Use Case Model	55
2.3	Mathematical Model	57
2.3.1	Relative Risk of Non-Diet Risk Factors	58
2.3.2	Prevalence of Smoking, Obesity, and Physical Inactivity	63
2.3.3	Cumulative Effect of Exposure to Diet-Related Risk Factors	67
2.3.4	Simple Formula for the Probability of Developing a Chronic Disease	71
2.3.5	Diet-Related Probability of Developing Cardiovascular Disease.....	73
2.3.6	Effect of Smoking, Obesity and Physical Inactivity	75
2.3.7	Differential Probabilities for Developing Cardiovascular Diseases	77
2.4	Analysis Class Diagrams	78
2.5	Web Page Architecture and Navigation.....	84
2.6	Data Model.....	85
2.6.1	Entity Relationships	87
2.6.2	Database Design.....	88
2.7	System Architecture	97
2.8	Subsystem Decomposition.....	99
2.8.1	Risk Calculator.....	99

2.8.2	Pattern Generator	101
2.8.3	General Information.....	103
2.8.4	User Management	104
2.9	Design Class Diagrams	104
3.	Implementation and Risk Assessment Results	111
3.1	Java Server Faces Technology Overview	111
3.1.1	JSF Component Tree.....	113
3.1.2	JSF Lifecycle	113
3.1.3	Java Studio Creator 2	115
3.2	Apache Tomcat Application Server.....	117
3.2.1	Tomcat Installation and Directory Structure.....	118
3.2.2	Class Loading.....	119
3.2.3	Database Connectivity	120
3.3	MySQL Relational Database Management System.....	120
3.4	Application Directory Structure.....	122
3.5	Graphical User Interface Implementation.....	125
3.5.1	Web Page Set Overview	125
3.5.2	Applet/JSF Integration	133
3.5.3	Food Intake Pattern Display.....	134
3.5.4	Images	138
3.6	Java Beans.....	138
3.7	System Specifications	142

3.8	Application Deployment.....	143
3.8.1	Generation of Web Archive File.....	144
3.8.2	Application Listener Class Configuration	145
3.8.3	Standard Libraries Installation.....	146
3.8.4	Database Connection Configuration.....	146
3.9	Food Intake Pattern Generation	149
3.9.1	Raw Data Preparation	150
3.9.2	Pattern Generation Script.....	151
3.9.3	Pattern Data Structure	151
3.9.4	Food Intake Pattern Template.....	157
3.10	Diet Modeling	159
3.10.1	Traditional American Diet.....	159
3.10.2	Mediterranean Diet	160
3.10.3	Vegetarian Diet.....	163
3.10.4	Vegan Diet	163
3.11	Reverse Risk Assessment Results.....	165
3.12	Conclusion	171
Appendix A. Use Case Sequence Diagrams.....		174
A.1	Use Case: Login.....	174
A.2	Use Case: Update Profile	175
A.3	Use Case: Look up Calendar.....	176
A.4	Use Case: Calculate Risk	177

A.5	Use Case: Browse Results	178
A.6	Use Case: Explore Pattern	179
A.7	Use Case: Look up Glossary	180
A.8	Use Case: Read General Information	181
A.9	Use Case: Use External Link	182
A.10	Use Case: Create New User Profile	183
Appendix B. Food Intake Pattern DTD		184
Appendix C. XSL Stylesheet Code.....		187
References.....		198

List of Figures

Figure 1: Schematic Representation of Direct Risk Assessment Process.....	5
Figure 2: Schematic Representation of Reverse Risk Assessment Process.....	6
Figure 3: Total Costs (in Billions of Dollars) of Major Cardiovascular Diseases.....	10
Figure 4: Classification of Chronic Disease Risk Factors.	13
Figure 5: Prevalence of Certain Health Conditions Among Adults (Ages 20-74) in the United States in 2000-2001 as Percentage of Total Population.....	14
Figure 6: Vitamin Classification.....	21
Figure 7: Essential Minerals.	26
Figure 8: Leading Causes of Death in the United States [13].....	33
Figure 9: Interdependence of Major Chronic Diseases and Risk Factors.....	34
Figure 10: Basic Risk Assessment Use Case Diagram.....	56
Figure 11: System Administrator Use Case Diagram.....	57
Figure 12: Collaboration Diagram for “Login” Use Case.	78
Figure 13: Collaboration Diagram for “Update Profile” Use Case.	79
Figure 14: Collaboration Diagram for “Look up Calendar” Use Case.....	79
Figure 15: Collaboration Diagram for “Calculate Risk” Use Case.	80
Figure 16: Collaboration Diagram for “Browse Results” Use Case.....	80
Figure 17: Collaboration Diagram for “Explore Pattern” Use Case.....	81
Figure 18: Collaboration Diagram for “Look up Glossary” Use Case.	82
Figure 19: Collaboration Diagram for “Look up Calendar” Use Case.....	82

Figure 20: Collaboration Diagram for “Use External Link” Use Case.....	83
Figure 21: Collaboration Diagram for “Create New Profile” Use Case.	83
Figure 22: Web Page Architecture and Navigation.	86
Figure 23: Diet-Related Risk Factor–Chronic Disease ERD.....	88
Figure 24: Food Intake Pattern ERD.	89
Figure 25: Bridge ERD (Food Intake Pattern and Risk Factor Entities).	90
Figure 26: Nutrient Class Structure of Food Intake Patterns.	91
Figure 27: The “Pattern” Table Schema.	92
Figure 28: The “User” Table Schema.	93
Figure 29: The Folder Structure of the XML Database.	94
Figure 30: Food Intake Pattern XML File Layout.	95
Figure 31: Reverse Risk Assessment System Architecture.	98
Figure 32: Reverse Risk Assessment Subsystem Decomposition.	100
Figure 33: Food Intake Pattern Generation Flowchart.....	102
Figure 34: Risk Calculator Design Class Diagram.	105
Figure 35: Login Design Class Diagram.	106
Figure 36: Update Profile Design Class Diagram.....	107
Figure 37: New User Design Class Diagram.	108
Figure 38: Simple Navigation Design Class Diagram.	109
Figure 39: Application Folder Hierarchy.....	123
Figure 40: The Application Package Class Files.	123
Figure 41: External Libraries Folder.....	124

Figure 42: Login Page.....	127
Figure 43: Home Page.	127
Figure 44: Informational Page - Cholesterol.	128
Figure 45: Informational Page - Fat.....	128
Figure 46: Informational Page - Heart Disease.....	129
Figure 47: Informational Page - Stroke.	129
Figure 48: Profile Page.	130
Figure 49: New User Page.	130
Figure 50: Calendar Page.....	131
Figure 51: Glossary Page.....	131
Figure 52: Risk Calculator Page.	132
Figure 53: Search Results Page.....	132
Figure 54: Top Part of Food Intake Pattern Page.	135
Figure 55: Energy-Yielding Nutrients Table - Food Intake Pattern Page.	136
Figure 56: Water-Soluble Vitamins Table - Food Intake Pattern Page.	136
Figure 57: Fat-Soluble Vitamins Table - Food Intake Pattern Page.....	137
Figure 58: Microminerals Table - Food Intake Pattern Page.....	137
Figure 59: Schematic Representation of Food Intake Pattern XML Structure.....	158
Figure 60: Risk of Cardiovascular Disease and Distribution of Vegan and Mediterranean Diet Patterns.....	166
Figure 61: Risk of Cardiovascular Disease and Distribution of Vegetarian and Traditional American Diet Patterns.	167

Figure 62: Average Diet-Related Risk of Contracting a Cardiovascular Disease.	169
Figure 63: Login Use Case Sequence Diagram.	174
Figure 64: Update Profile Use Case Sequence Diagram.	175
Figure 65: Look up Glossary Use Case Sequence Diagram.	176
Figure 66: Calculate Risk Use Case Sequence Diagram.	177
Figure 67: Browse Results Use Case Sequence Diagram.	178
Figure 68: Explore Pattern Use Case Sequence Diagram.	179
Figure 69: Look up Glossary Use Case Sequence Diagram.	180
Figure 70: Read General Info Use Case Sequence Diagram.	181
Figure 71: Use External Link Use Case Sequence Diagram.	182
Figure 72: Create New User Profile Use Case Sequence Diagram.	183

List of Tables

Table 1: System Requirements	52
Table 2: Prevalence of Cardiovascular Disease among Different Ethnic Groups.	62
Table 3: Prevalence of Smoking, Obesity and Physical Inactivity.	64
Table 4: Average Daily Consumption of Hazardous Substances.	74
Table 5: Average Daily Exposure to Non-Diet Hazards and CVD Probability.....	76
Table 6: Prevalence of Cardiovascular Diseases (2003).....	77
Table 7: Applet Definition and Configuration in JSP.....	134
Table 8: Listing of Application Package Java Classes and JSP.....	139
Table 9: System Hardware Specifications.	143
Table 10: System Software Specifications.....	143
Table 11: Listener Class Configuration Code.....	145
Table 12: Database Connection Configuration.....	147
Table 13: Food Intake Pattern Data Structure Code.	151
Table 14: Diet Type Definitions.	162
Table 15: Prevalence of Heart Disease and Hypertension: Comparison of Results Obtained with Statistical Data.	169

1. Reverse Risk Assessment and Diet-Related Chronic Diseases

The focus of this chapter is on the analysis of the reverse risk assessment problem as applied to risk evaluation of contracting diet-related chronic diseases.

1.1 Introduction

The process of risk assessment plays an important role in many industries including insurance, investment and health care to name a few. Risk assessment is a part of the broader concept of risk management which gained popularity and became an established term about 30 years ago, when global markets experienced high levels of volatility [1]. The definition of risk as exposure to uncertainty may seem simple yet it has a profound meaning. In particular, the notion of risk has two components: (1) uncertainty and (2) exposure to it [2-5]. The uncertainty component is highly subjective. What seems uncertain to one person may be crystal clear to other people. Uncertainty is removed when more information becomes available to the exposed person about an object, event or circumstances associated with a risky situation. Sometimes, risk is paralleled with ignorance or disregard of important information as related to the degree of uncertainty.

Generally, risk can be viewed as a collection of pathways in time and configuration space with different consequences. Those pathways may or may not

realize. In other words, there is a finite probability that each pathway will occur. Thus, quantitatively, risk means an expected value resulting from a sum of products of possible pathways and their probabilities. A risky situation corresponds to a set of pathways with essentially different consequences but with more or less comparable probabilities for each pathway realization. Hence, a situation is no longer risky if there is only one scenario with a probability close to unity. In the latter case, the uncertainty is almost zero and we know for sure that there will be only one outcome. The subjectivity of risk perception is closely related to a personal assessment of the realization probability for a particular event sequence.

As human beings, we strive to reduce uncertainty and get more information about possible development of events. Reduction in uncertainty is crucial for mobilizing available resources and developing the most efficient behavioral pattern to handle environmental changes. A particular situation may become extremely worrisome if there is no apparent way to reduce uncertainty. Circumstances of this kind exhaust our emotional, psychological and physical resources. Therefore, proper risk management is of crucial importance for survival of every living creature.

This thesis focuses on the subject of the so-called reverse risk assessment. Specifically, it considers an issue of risk evaluation for developing diet-related chronic diseases. In general, a solution to a reverse risk assessment problem is provided by generation of a large number of possible scenarios which lead to a given probability of the undesired outcome. In the case of diet-related diseases, different risky scenarios correspond to various daily food intake patterns. For this purpose, a software application

has been designed and implemented to generate various food intake patterns which could be searched and examined for potential risk mitigation.

1.2 Direct versus Reverse Risk Assessment

Risk assessment is a process of examination of potential hazards or exposure to unfavorable conditions, which may result in a certain kind of harm, damage or loss [6].

Thus, risk assessment implies four main tasks:

- (A) Analysis of objects, processes, people or other subjects of interest, which are likely to be jeopardized by unfavorable changing conditions, or undesirable events in the future,
- (B) Determination of risk factors that may contribute to risk of a certain event to occur,
- (C) Evaluation of undesirable events and their impact on subjects or objects on interests,
- (D) Estimation of the probability of the undesirable event(s) or unfavorable change in existing conditions.

The results of the risk assessment process are used later as a basis for risk management and mitigation.

Most people are familiar with the direct risk assessment, or simply known as risk assessment. In direct risk assessment problem, activities (A), (B) and (C) are done first and then the probability of a certain undesirable event is estimated (D). Item (A) is a

starting point of the risk assessment process and essentially implies analysis of the existing or predefined initial conditions, states or processes, whose stability may be threatened in the future. Activities (B) and (C) deal with specification and evaluation of risk-bearing events and conditions. Item (D) is related to quantification of unexpected event occurrence and, in many instances, is closely-linked to the risk impact evaluation. Few examples considered below illustrate the direct and reverse risk assessments concepts.

One of the widely-publicized direct risk assessment processes was the attempt to deal with hazardous agents in 1950s. Toxicologists examined existing data of various population groups and hazardous chemicals, such as pesticides and food additives. They evaluated impact of exposure to those chemicals on human health and estimated probability of toxicity risks for population. The end-result of this effort was a publication of a Food and Drug Administration (FDA) report which defined the basis for what is now called the acceptable daily intake (ADI) [7, 8]. Technically, ADI represents a threshold intake of a chemical for a very large population of people, below which toxicologists don't expect significant toxicity risks. The idea of threshold takes its roots from the natural ability of the human body to neutralize the effect of toxic substances and recover from intoxication for small enough doses. In reality, there is always a certain probability of intoxication due to chemical exposure slightly below and above the ADI level. As a rule, the intoxication probability is a smooth function of exposure to hazardous substances (especially those which are both toxic and carcinogenic) and the ability of the human immune system to fight intoxication varies from one individual to another.

Therefore, the intoxication threshold may not exist in nature. However, ADI plays an important role of a convenient reference point, below which toxicity risks are thought to be acceptable for a large and variable population group.

The aforesaid example clearly shows all four (A, B, C and D) stages of the direct risk assessment process. For the purpose of the quantitative risk assessment, it is important to identify major entities involved in the direct risk evaluation process. They are schematically shown in Fig. 1.

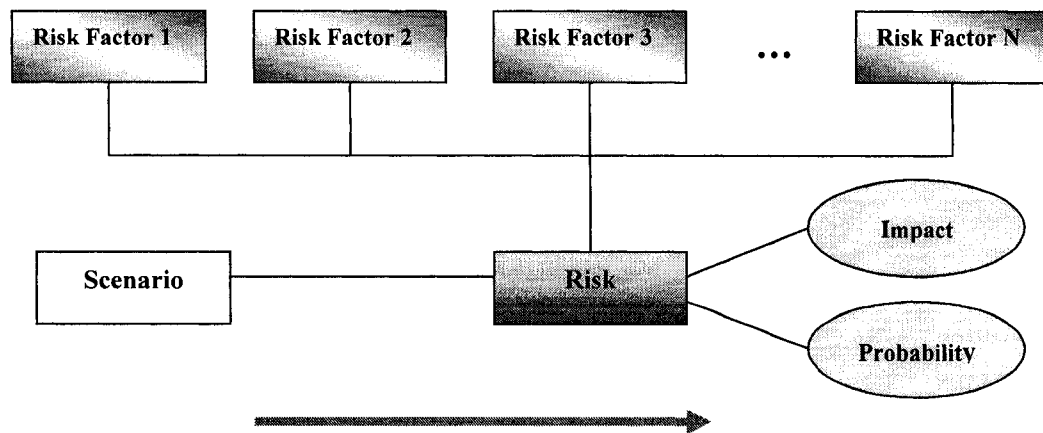


Figure 1: Schematic Representation of Direct Risk Assessment Process.

Reverse risk assessment implies an opposite process. First, one has to determine how much risk in terms of probability for a certain unfavorable event can be tolerated. Secondly, the impact of the event is estimated. Thirdly, risk factors associated with the

undesirable event, as well as possible contributing environmental conditions are evaluated and analyzed. Finally, a variety of possible initial scenarios, which may result in materialization of the undesirable-event probability, given the risk factors considered, are calculated and described. Although the process of reverse risk assessment looks like a simple reversal of the direct risk assessment (D->C->B->A), there is a fundamental difference between the two approaches. In the direct risk assessment, there is only one well-defined set of initial conditions which, normally, exist in a real-world situation. After being exposed to the specified risk factors, this set of conditions may lead to a certain unfavorable development with a certain probability value. In contrast to that, the

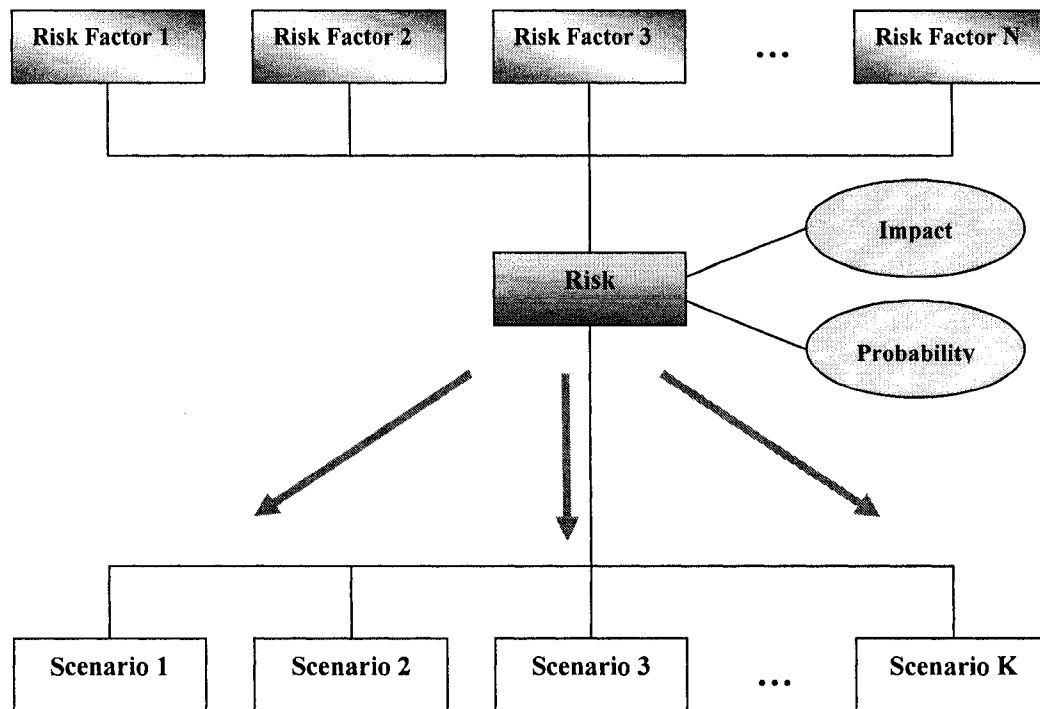


Figure 2: Schematic Representation of Reverse Risk Assessment Process

reverse risk assessment attempts to find all possible initial conditions and scenarios, which may or may not exist in reality and which satisfy given risk probability value and combination of risk factors. It should be stressed that the number of potential situations in a reverse risk assessment problem may be huge as opposed to one well-specified initial state in the case of direct risk assessment. The reverse risk assessment process is illustrated by Fig. 2.

Examples of reverse risk assessment are plenty. For instance, they include (a) evaluation of personal goals and potential scenarios awaiting a person in the future if no attempts are undertaken to change his/her present position or lifestyle, (b) notifications and implications for symptom-free donors when a blood recipient contracts a blood-transfer related disease, (c) estimation of cleanup goals and sewage treatment objectives by organizations and facilities which contaminate the physical environment [9, 10], (d) reduction in the amount of animal testing to meet a given set of information requirements [11]. Some of those examples deserve few additional comments. For instance, example (a) actually considers the present situation and a variety of potential future scenarios, which may develop, if, with the 100% probability, the decision-maker doesn't undertake any action to change his/her status quo. In example (b), various implications are considered for presumably risk-free donors under the assumption that with a certain probability a recipient contracts a blood-related disease. Similarly, a certain level of environment contamination (example (c)) implies various clean up measures to be undertaken by the polluting facilities and so forth. All these examples emphasize one of the major values of the reverse risk assessment findings - making recommendations on

how to modify the existing conditions, processes, etc. in order to reduce or otherwise eliminate potential risk.

In summary, the reverse risk assessment problem deals with situations when risk is known or predefined and the attempt is made to predict, describe or otherwise provide a variety of conditions, which are characterized by the specified degree of risk. Although being similar in many aspects to the usual risk assessment methods, the reverse risk assessment is run in reverse. The reverse risk assessment normally implies calculating various scenarios which may cause the specified risk. Typically, several risk factors may contribute to the same risk outcome and the reverse risk assessment objective is to find all potential combinations of the associated conditions with the same cumulative effect. Needless to say, that even a relatively simple risk evaluation problem can involve a huge number of potential scenarios. Since, the reverse risk assessment is crucial for decision-making process, it is of paramount importance to facilitate generation and analysis of risk-associated conditions. Such a task most efficiently can be done by a software system designed to account for numerous constraints and contributing risk factors.

1.3 Problem of Reverse Risk Assessment and Diet-Related Chronic Diseases

It is now widely recognized that apart from the so-called non-modifiable risks like gender, ethnicity, and genetic predisposition, diet is one of the crucial lifestyle factors of our choice, which influence development of chronic diseases. Dietary risk factors are

directly linked to the major chronic diseases, like atherosclerosis, stroke and hypertension, which pose the greatest threat to the lives of people in developed countries. This kind of diseases develops over a lifetime as a result of metabolic abnormalities induced by variety of factors including diet. Diseases and ill health can not only ruin people lives but also greatly affect businesses and organizations in terms of productivity and related insurance costs.

Since food intake affects the human body metabolism on day-to-day basis, diet significantly contributes to development of variety of chronic diseases including (a) cardiovascular disease, (b) hypertension, (c) cancer, (d) strokes, (e) diabetes, (f) osteoporosis, (g) liver and (h) kidney disorders, to mention a few. In particular, cardiovascular disease, cancer and stroke account for two thirds of the nation's two million deaths every year [12]. Impact of chronic diseases on the American society is enormous. For instance, medical costs of treatment heart disease, cancer, stroke and diabetes, which are attributed to diet, amount to \$33 billion each year. It is estimated that those chronic diseases result in \$9 billion of lost productivity annually [12]. Figure 3 shows estimated direct and indirect costs in billions of dollars, associated with cardiovascular diseases and stroke in the United States. The data presented include direct health expenditures, as well as indirect costs of lost productivity due to morbidity and mortality [12].

Currently, it is believed that a balance diet is a key factor in prevention of chronic diseases. Such a diet is supposed to provide about 40 essential nutrients in the amounts

which are healthy for a human body. Lack or excess of these nutrients would pose a health risk. In addition to that, there are at least 5 traditional diet components such as

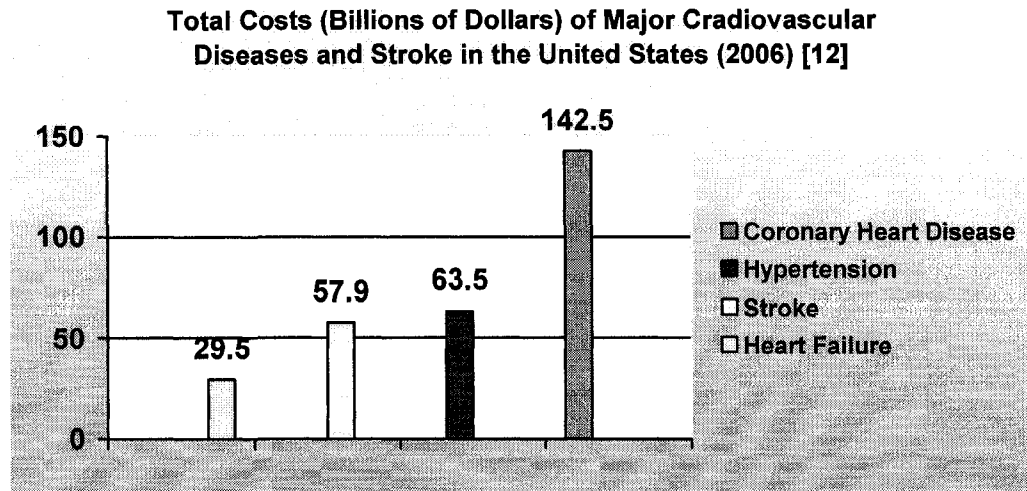


Figure 3: Total Costs (in Billions of Dollars) of Major Cardiovascular Diseases.

cholesterol, refined sugar, table salt, alcohol, and trans fat which are not required for healthy metabolism, but may significantly contribute to the risk of contracting a chronic disease. The large set of risk factors, which influence development of diet-related chronic diseases, poses a challenge and, at the same time, represents an opportunity for reverse risk assessment. Indeed, different combinations of environmental factors and diets can result in the same probability of contracting a chronic disease. Risk assessment, as applied to dietary risk factors, deals with identifying major risks associated with food intake patterns and taking precautions, which can prevent harm to human health. Thus, the reverse risk assessment objective is to find a variety of possible diets and

environmental factor combinations, which may result in a chronic disease within a specified period of time with a certain probability.

Modern Americans consume a large variety of foods, including vegetables, fruits, meats, dairy products, cereals, spices etc. The food databanks typically contain thousands of uniquely-different food entries. Food combinations and ingested amounts may vary considerably on a daily basis. Therefore, from the user perspective, the reverse risk assessment enormously facilitate decision-taking process, since it is much easier to perform a simple search on the large set of computed-in-advance scenarios rather than to go through a lengthy process of specifying a list of potentially 40-50 food items, which would provide adequate daily intakes of about 40 essential nutrients.

The next section takes a closer look at chronic disease risk factors to be accounted for in the reverse risk assessment problem.

1.4 Disease Risk Factors

Disease risk factors are defined as traits and lifestyle habits that influence the probability of contracting a certain disease, but are not necessarily causal [13]. Hence, all risk factors can be divided into two groups. One includes the trait risk factors such as gender, ethnicity, heredity, unhealthy reaction to stress, and age-associated hormonal changes in the body. These risk factors cannot be controlled. For instance, men are more susceptible to cardiovascular disease and develop it at an early age than women. Atherosclerosis

usually doesn't affect women in terms of vascular disease until after menopause. Thus, hormonal changes, which are beyond control of any individual, may play an important role in contracting a certain condition.

The other group comprises controllable, or habit risk factors. Typical controllable risks include smoking and dietary risk factors. Classification of various risk factors is presented in Fig. 4. Also shown in controllable risk listing are conditions and diseases acquired as a result of exposure to the primary risk factors and unhealthy habits. Generally speaking, almost every chronic disease is a risk factor for developing another chronic condition, since one disease often serves as a precursor for contracting another. Only the first top three conditions are indicated on Fig.4 for the sake of clarity. Those include high blood cholesterol, high blood pressure, and obesity.

Among 21 risk factors shown in Fig.4, diet-related risks are the ones which vary greatly on the day-to-day basis. Variations in exposure to those risks are associated with the variety of foods and their amounts an individual consumes with his/her meals. The subset of controllable risk factors such as obesity, high blood cholesterol and high blood pressure deserves few additional comments. These risk factors are considered to be secondary, since, in the overwhelming majority of cases, they are developed as a result of exposure to other risk factors. In particular, the spread of obesity is strongly correlated with the availability of food and reduced need for physical labor in the developed countries. By the same token, high sodium diet and high fat intake combined with consumption of animal foods rich in cholesterol are major factors which increase blood cholesterol and contributes to hypertension.

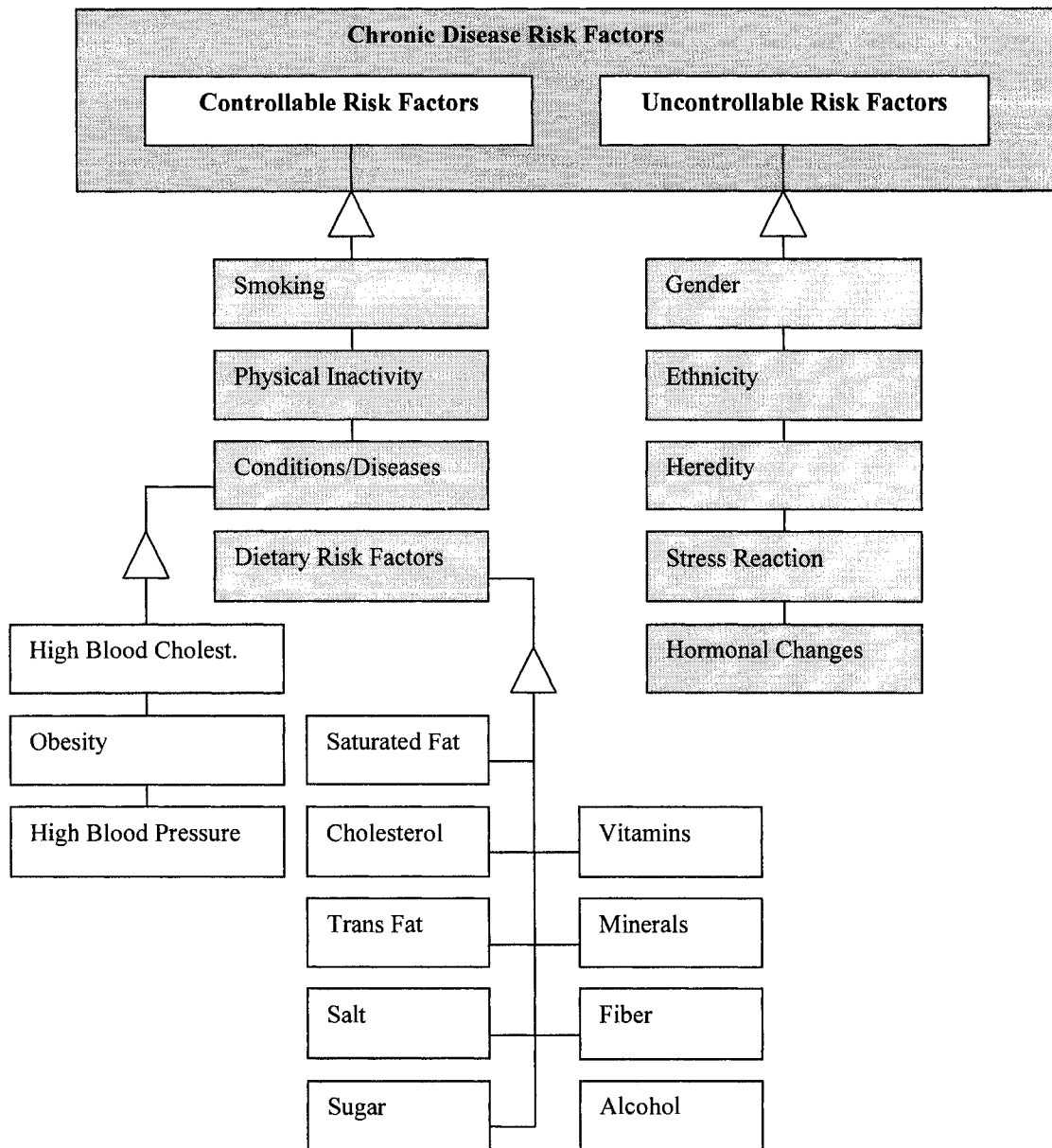


Figure 4: Classification of Chronic Disease Risk Factors.

Fig. 5 displays the prevalence of certain health conditions as risk factors for chronic diseases. Also shown for comparison are primary controllable risk factors – smoking and physical inactivity. It should be noted that Fig. 5 shows the lowest estimate for the

Prevalence of Health Conditions as Risk Factors for Chronic Diseases Among U.S. Adults (% Total Population) [12,14]

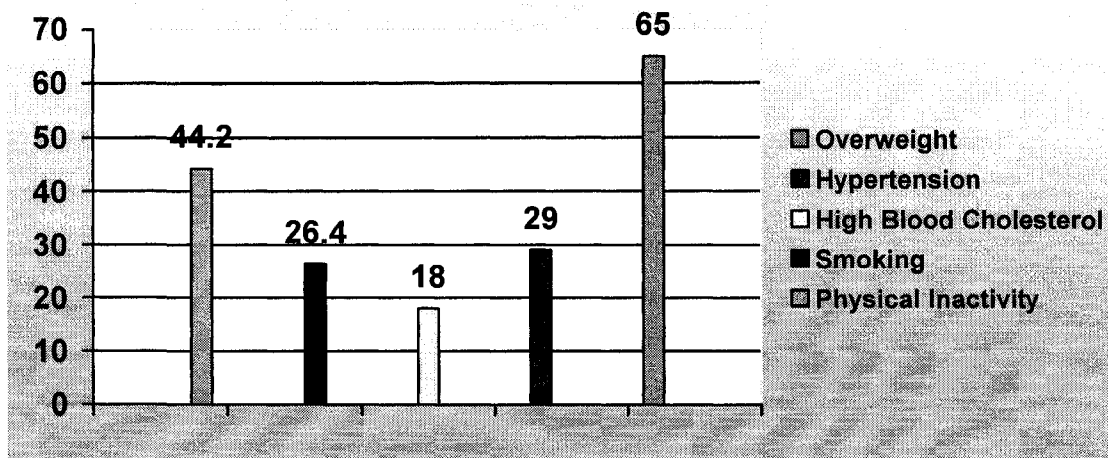


Figure 5: Prevalence of Certain Health Conditions Among Adults (Ages 20-74) in the United States in 2000-2001 as Percentage of Total Population.

percentage of physically inactive Americans. The real number may be higher due to subjective nature of this estimate [14]. Since our focus is primarily on the dietary habits, the next section discusses the respective risk factors in more detail.

1.5 Diet-Related Risk Factors

As it follows from the previous section, diet-related risk factors influence not only development of chronic diseases but also may trigger creation of secondary risk factors like being overweight or obese. This section provides a detailed overview of major dietary risk factors, which are associated with chronic diseases.

1.5.1 Fat

Fat is a subset of the class of nutrients known as lipids. The lipid family includes triglycerides (fats and oils), phospholipids, and sterols. Triglycerides provide the body with energy and protect it from cold and mechanical shock. Phospholipids and sterols contribute mostly to cell structure and serve as a raw material for some hormones, vitamin D and bile [13].

Triglycerides consist of fatty acids and glycerol. Individual fatty acids may have different physiological effect on human body, which depends on a specific mix of fatty acids in the diet. One of the important characteristics of fatty acids is the degree of saturation which is defined as a number of double bonds in a triglyceride molecule. Thus, fatty acids are classified as (a) saturated (no double bonds), (b) monounsaturated (one double bond), and (c) polyunsaturated (more than one double bond). Among polyunsaturated fatty acids, there are two acids, which are not synthesized in the body and play a prominent role in human health. One of them is called linoleic (belongs to the group of omega-6 fatty acids). Another one is the so-called α -linolenic (one of the omega-3 fatty acids). A lack of dietary linoleic acid causes rough, scaly skin and dermatitis. Primary sources of this acid are liquid canola oil, soybean oil, corn oil, and safflower oil [13]. A deficiency of α -linolenic acid in the diet results in scaly and hemorrhagic dermatitis, hemorrhagic folliculitis of the scalp, impaired wound healing, and growth retardation. It is obtained from plant sources including soybean oil, canola oil, walnuts, and flaxseed [13].

Unsaturated fatty acids have an important isomeric variation, which is called trans fatty acids. The difference between trans fatty acids and normal unsaturated fatty acids is in configuration of the double bond. In the case of trans fatty acids, the double bond is characterized by a trans configuration as opposed to the normal cis-configuration. The Latin word cis means that nearest hydrogen atoms are oriented in the same direction with respect to the double bond linking two carbon atoms.

Relatively recently scientists discovered that trans fatty acids have a detrimental effect on human health and noticeably affect the blood lipid profile. Trans fatty acids are mostly derived as a result of hydrogenation of polyunsaturated oils. Increased trans fatty acid content hardens the oil and prevents it from being spoiled. Typical sources of trans fatty acids include shortening and commercially baked goods, snack foods, fried foods and margarine [15]. Current recommendation regarding trans fat intake is to avoid it as much as possible.

Lipids are transported in the blood by four types of particles: chylomicrons, VLDL (very low density lipoproteins), LDL (low density lipoproteins) and HDL (high density lipoproteins). Concentration of these particles in the blood constitutes the blood lipid profile, which is a good indicator of a healthy fat intake. A desirable blood lipid profile is characterized by four major components: (1) total cholesterol < 200 mg/dl (5.2 mmol/L), (2) LDL cholesterol < 130 mg/dL (3.4 mmol/L), (3) HDL cholesterol > 35 mg/dL (0.9 mmol/L), (4) triglycerides < 200 mg/dL (2.3 mmol/L). It should be noted that some researchers consider the level of cholesterol of 150 mg/dl or less to be safe. Indeed,

about 100 million Americans have elevated level of cholesterol (above 200 mg/dl) with the average cholesterol level of heart disease victims at 225 mg/dl [15].

High fat intake is believed to be one of the major flaws of westernized diet, contributing to heart disease, obesity, development of cancer and other health problems. Saturated fat intake has been found to linearly increase blood LDL, thus increasing the risk of cardiovascular disease [15]. Quantitative studies show that the two major components of high fat intake, saturated fat and trans fat, contribute differently to the risk of cardiovascular disease. In particular, there is a positive dose-response relationship between saturated fat intake and the LDL level in the blood. Each 2 g intake of saturated fat increases the total LDL by 1.5 mg/dl, which is equivalent of 1.5% increase in the risk of cardiovascular disease. Similarly, every 1 g intake of trans fat increases the risk of cardiovascular disease, on average, by at least 4.5% [15].

It has been proved that a diet low in both fat and saturated fat, combined with regular fish consumption, produces an optimal blood lipid profile [15-17]. The National Advisory Committee on Nutritional Education and the Committee on Diet and Health recommend the following guidelines for total and saturated fat intakes: (1) total fat intake should be less than 30% of energy intake (currently 35%), (2) saturated fat intake must be less than 7 % of energy intake (currently 12%), and (3) trans fat intake must be less than 1% of energy intake (currently 2.6%) [15].

1.5.2 Cholesterol

Cholesterol is a soft waxy substance which belongs to the group of lipids called sterols. It is common in all animal tissues. Cholesterol is manufactured in the liver and is essential

for normal body functions including the formation of cell membranes, production of hormones, bile acid, and vitamin D. Human body is able to synthesize sufficient amount of cholesterol to meet biologic requirements. At present, there is no evidence for a dietary requirement for cholesterol [15].

Serum cholesterol is transported in the blood mainly by LDL (60%-70% of total cholesterol) and HDL (20%-30%). Elevated level of LDL correlates positively with increased risk of cardiovascular diseases. Dietary intake of cholesterol increases the overall concentration of LDL in the blood, thereby increasing the risk of heart disease and other chronic conditions. More specifically, cholesterol intake increases the total serum cholesterol at the rate of 5 mg/dl per each 100 mg of added dietary cholesterol in the intake range from 0 to 400 mg [15]. The relationship between cholesterol intake and change of the cholesterol blood concentration is almost linear in the aforesaid intake interval. However, at the intake of above 300 mg the dependence becomes weaker and every 100 mg of dietary cholesterol results in 1.5 mg/dl increase of the serum cholesterol. Analytical approximations of the serum cholesterol concentration change as a function of cholesterol intake predict, on average, from 2 to 3 mg/dl in concentration change for every 100 mg intake of dietary cholesterol. For a person with the serum cholesterol level of about 180-200 mg/dl, this change would result in the LDL level increase of about 1%. Based on the current statistical evidence, it is expected that every 1% of extra LDL will increase the risk of coronary heart disease by 1% [15]. Therefore, the intake of dietary cholesterol should be kept as minimal as possible.

1.5.3 Fiber

Fibers represent a separate group of complex carbohydrates and are the structural parts of plants. They can be found in all plant-derived foods such as grains, fruits, vegetables and legumes. Fibers possess unique physical properties which make them indispensable food components from the digestive system point of view. Indeed, fibers are able to capture water and, in this sense, they work like a sponge, increasing the bulk of intestine content. They form a viscous, gel-like solution, which is easily transported by bowel movements. Fibers can be fermented – broken down by bacteria into digestible fragments. Finally, fibers bind bile and minerals.

The FDA sets a daily value for fiber at 25 g per 2200-kcalorie intake. The American Dietetic Association suggests from 20 to 35 grams of fiber depending on energy needs [18]. Currently, Americans consume about 12–13 g of fiber a day, which is twice as less fiber as recommended for a healthy digestion process.

1.5.4 Vitamins

Vitamins are chemically-active substances, which differ from the carbohydrates, fats and proteins in a variety of ways. Their molecules are relatively short and not linked into long chains. Vitamins don't yield usable energy when broken down. Instead, they assist the enzymes that release energy from carbohydrates, fats and proteins. The amount of vitamins needed on a daily basis is small and measured in milligrams or micrograms.

Vitamin deficiency can cause blindness, bone growth retardation, mental confusion and increase the overall disease risk. In the United States, people spend billions

of dollars each year on vitamin supplements, in an attempt to ensure themselves against specific diseases.

All vitamins are conventionally divided into two groups depending on their solubility properties [13]. The hydrophilic or water-soluble vitamins include all B-vitamins and vitamin C. Vitamins A, D, E and K are hydrophobic or fat-soluble. Although the vitamins are essential nutrients, they may be harmful if taken in large quantities. Therefore, both inadequate intake and over-consumption of vitamins are dietary risk factors. Vitamin toxicity levels are most likely reached when vitamins are consumed from supplements. Fig. 6 provides a high level view of vitamin classification. Vitamins functions and availability in foods is briefly discussed below.

Vitamin A, also known as retinol, stimulates and maintains healthy skin, growth of bones, resistance to infection and night vision. This vitamin is found in carrots, spinach, peppers, butter, margarine, watercress salad, dried apricots, full-fat dairy products like sour cream, yogurt and cheese. In plant foods vitamin A is mostly present as its precursor, beta-carotene.

Vitamin B1, or thiamin, participates in the process of releasing energy of carbohydrates. Among healthy foods containing thiamin are sunflower seeds, tuna, various kinds of beans and lentils, rice, bran, oatmeal, rye flour, whole grain bread, fortified corn flakes and wheat cereal.

Vitamin B2, or riboflavin, used to convert proteins, fats and carbohydrates into energy and for the growth and repair of tissues and healthy skin. It is present in spinach,

Romaine salad, spelt flour, almonds, yogurt, cheese, whole grain bread, dried prunes, mushrooms, cashews, millet, and avocados.

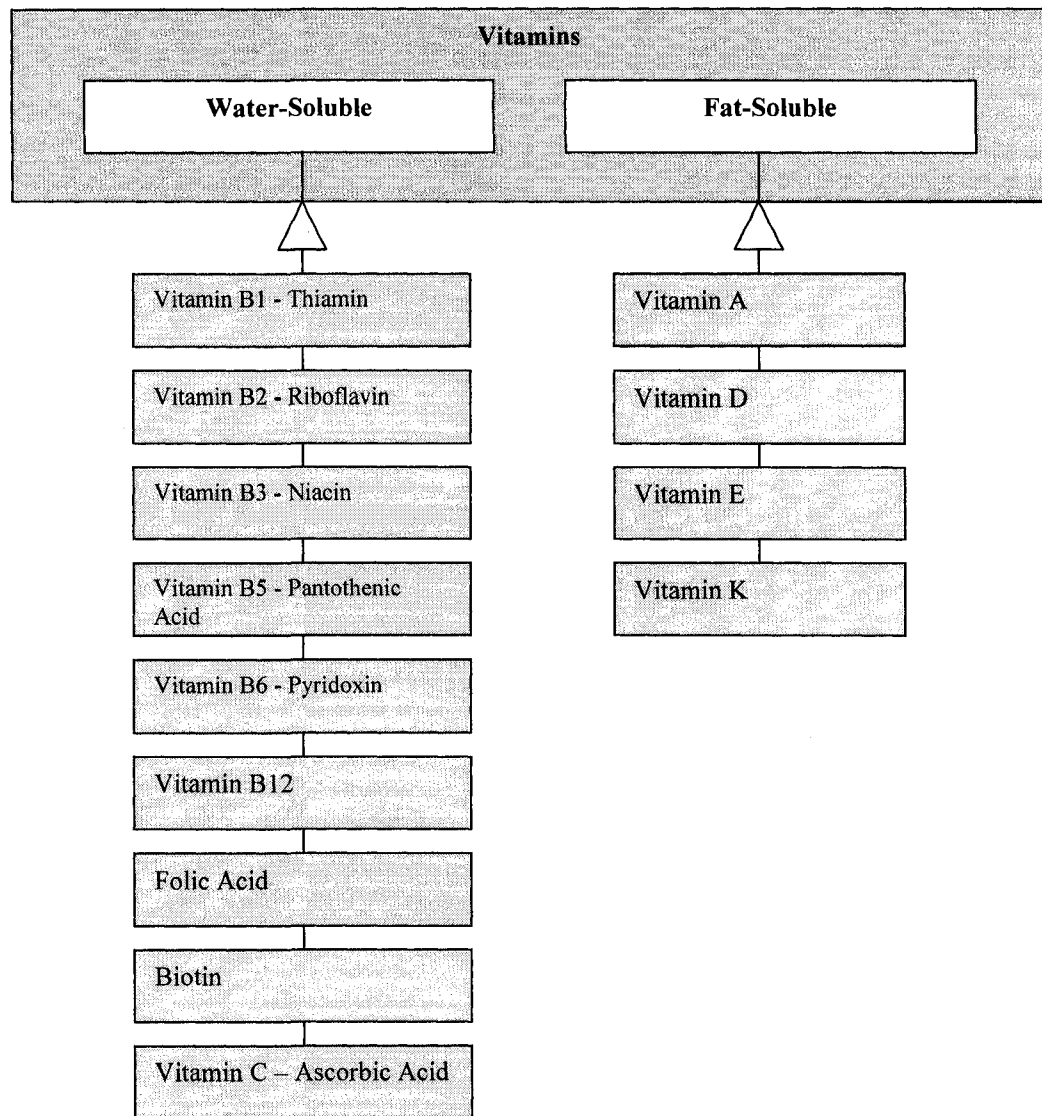


Figure 6: Vitamin Classification.

Vitamin B3, also called niacin, participates in the energy production cycle, is important for maintaining healthy skin, and the nervous system. Vitamin B3 plays an important role in fat metabolism and regulation of insulin activities. The vitamin is found in a variety of foods of animal origin like chicken, lamb, turkey, tuna and salmon. It is also present in significant amounts in spelt flour, peanuts, whole grain bread, mushrooms, and sesame seeds.

Vitamin B5, mostly known as Pantothenic Acid, plays an important role in energy production and antibody formation. The vitamin is widely found in most foods, including mushrooms, sunflower seeds, broccoli, low-fat yogurt, and strawberries.

Vitamin B6 (pyridoxin) facilitates red blood cell formation and protein metabolism. It is found in foods of animal origin, bran, whole grain bread, hazelnuts, bananas, peanuts, currants, potatoes. The deficiency of vitamin B6 results in skin disorders like eczema and seborrheic dermatitis, anemia, convulsions or seizures.

Vitamin B12 takes part in red blood cell formation, stimulates growth, and a healthy nervous system. Unlike most vitamins of the B-vitamin family, vitamin B12 cannot be produced by plants and animals. So far, only microorganisms like mold, bacteria, yeast and algae have been identified as exclusive original sources of this vitamin. Vitamin B12 is responsible for blood cell formation, protein metabolism and nerve cell development. The most efficient sources of vitamin B12 are dishes prepared from fish like mackerel and salmon. Other vitamin B12-rich foods include meats, poultry fortified soy milk and cereals.

Folic acid, or shortly folate, derives its name from French foliage, meaning the abundance of the vitamin in plant foods. Folate facilitates red blood cell formation, protein synthesis and DNA metabolism. Good sources of folate are spinach, broccoli, peanuts, almonds, hazelnuts, corn flakes, beans, and rice. Folate deficiency has been associated with increased risk of heart disease.

Biotin participates in energy production process, synthesis and breakdown of some fatty and amino acids. Biotin is found in a variety of foods, including egg yolks, leafy vegetables and nuts.

Vitamin C is probably one of the best known vitamins. It is a powerful antioxidant, helps to maintain healthy skin, bones, teeth and gums, enhances resistance to infection and wound healing, and participates in energy production and growth. Vitamin C is present in large quantities in citrus fruits, broccoli, spinach, berries, and peppers.

Vitamin D, also known as calciferol, stimulates absorption of calcium and phosphate and contributes to healthy bones and teeth. It is also responsible for maintaining a proper calcium concentration in the blood. The best sources of calciferol are salmon, shrimp, eggs and milk.

Vitamin E acts as an antioxidant protecting vitamins A and C and other important substances in the body. It is present in significant amounts in sunflower seeds, almonds, leafy vegetables, berries, vegetable oils, wheat germ, hazelnuts, avocados.

Vitamin K plays an important role in effective blood clotting. Found in spinach, cabbage, cauliflower. Vitamin K is also obtained from bacterial synthesis in the intestine. Low levels of vitamin K in the blood are associated with increased risk of fractures due to

calcium loss in elderly people, especially women. Excellent sources of vitamin K are leafy vegetables, like Swiss chard, spinach, kale and mustard greens. Other good sources include Brussels sprouts, cabbage, carrots and broccoli [13].

1.5.5 Minerals

The main minerals required in the diet are calcium, magnesium, sodium, potassium, chloride, phosphorus and sulfur. Other minerals which are necessary in only tiny quantities (less than 100mg/day) are called trace elements. The first group of minerals is often referred to as macro-minerals, since the typical adequate intake of macro-minerals is of the order of few milligrams. The second group is called micro-minerals and their typical adequate daily intake is in the range of micrograms. Fig. 7 shows the listing of major macro- and micro-minerals required for maintaining normal body functions.

Calcium, magnesium, and phosphorus form the structure of bones and teeth. Sodium, potassium and chloride are crucial for maintaining the body's fluid balance. Sulfur is an important component of special rigid proteins contained in skin hair and nails. Mineral bioavailability depends on food composition and may vary considerably. Fortunately, minerals are not destroyed by heat, air or acid and only little care is needed to preserve minerals during food preparation [13]. Minerals can be lost from food mainly when they leach into water, which is then thrown away.

Below we focus on few major minerals which are usually associated with well-defined risk of chronic diseases. Those include sodium, potassium and calcium. Both sodium and potassium play a prominent role in the body's water balance controlling the composition of blood and other body fluids. Sodium chloride (salt) is present in

processed foods and in small amounts in vegetables, fruits and grains. Most people consume too much sodium which can lead to high blood pressure. The impact of sodium intake on the blood pressure is so significant that table salt, as a primary supplier of sodium, is identified as one of the dietary risk factors (See section “Salt” in this chapter).

Potassium is widely found in plant foods, especially root vegetables and wholegrain cereals. Sodium deficiency is extremely rare, whereas lack of potassium can cause excessive vomiting or diarrhea.

Calcium is used in the body for building and maintaining strong bones and teeth, and is also responsible for muscle contraction and blood clotting. This element is found in dairy products, leafy green vegetables, almonds, sesame seeds, dried fruit, pulses, fortified soy milks. Low calcium concentration in extra-cellular fluids can result in the condition called tetany. Tetany is characterized by severe muscles contractions, spasms, and cramps. When the calcium concentration drops below normal, the nervous system becomes increasingly excitable, and nerve impulses are generated spontaneously, sending signals to skeletal muscles and causing spasmodic contractions, which often manifest in tingling in the fingers, toes, and lips. Almost 99% of body calcium, however, is in bones and teeth. Bone tissue is built by the process of crystallization of calcium salts in the protein matrix called collagen. The density of calcium-based crystalline formations determines the strength and rigidity of the bones. The mature bones are able to support weight of human beings by the time they learn how to walk [13]. The chemical composition of the bone tissue doesn't stay the same. Instead, bones are in a continuous

process of adjustment and remodeling. Healthy bones normally maintain a good balance in terms of losing and gaining calcium from food.

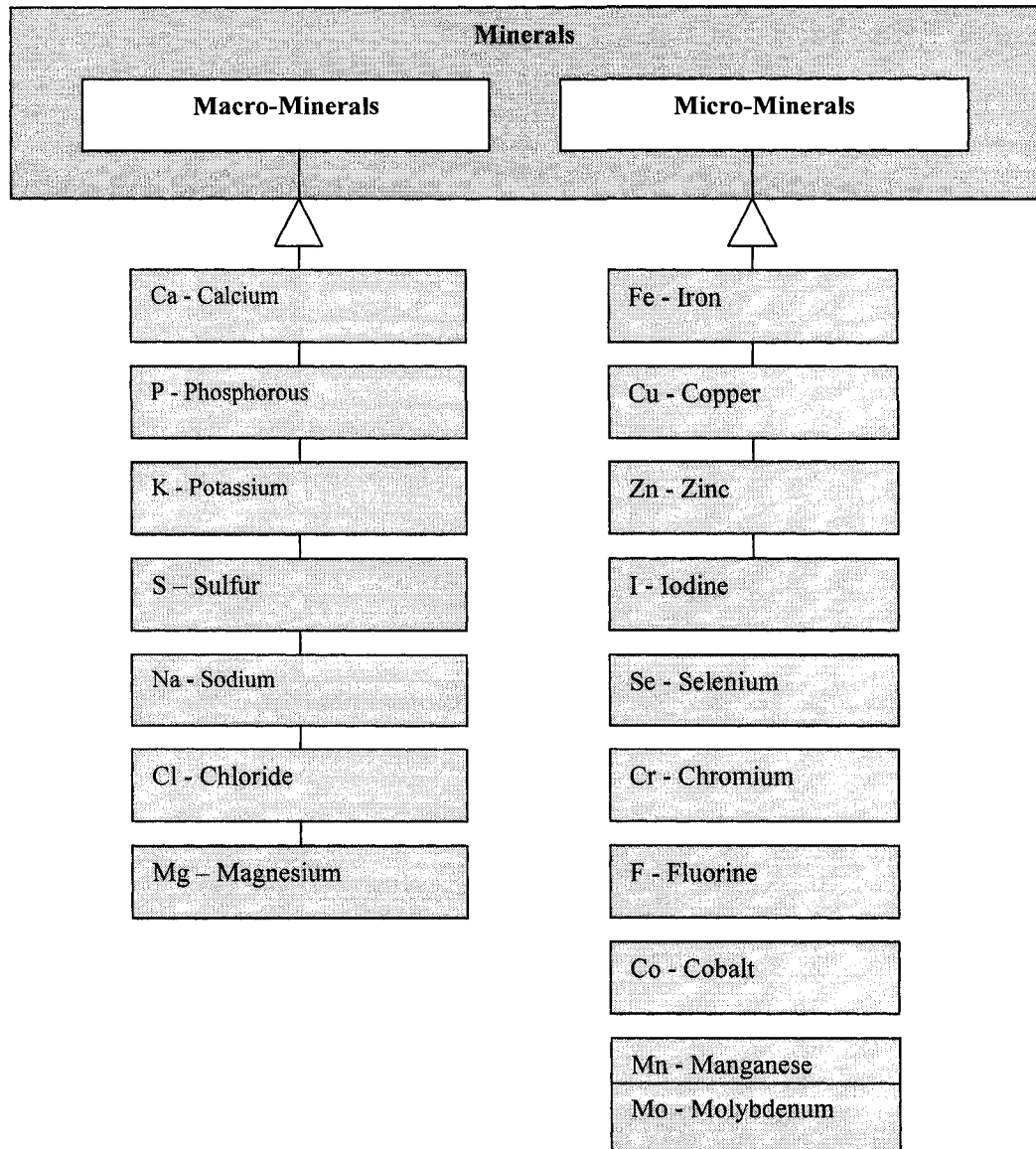


Figure 7: Essential Minerals.

Calcium absorption in the body is regulated by the calcium-binding protein, which is released when calcium concentration is too low. Normally, adult people absorb 30% of calcium ingested with food, but absorption rate for pregnant women increases up to 50% to provide enough building material to the body of the future child. Interestingly, some substances, like vitamin D, increase calcium absorption when their concentration is high and inhibit absorption of calcium when there is shortage of those substances. Excessive intake of fiber is known to inhibit calcium absorption. That is why vegetables and fruits are not efficient sources of calcium [13]. Relatively high availability and concentration of calcium is found in foods like canned sardines, milk, yogurt and cheese.

1.5.6 Sugar

If the impact of some of the dietary risk factors is still being debated, there is a remarkable consensus among scientists, nutritionist and doctors on the relationship between refined sugar intake and dental caries [13, 15, 19]. Added sugar directly provides fermentation material to mouth bacteria, which would otherwise produce sugar by breaking starches and fibers. The subsequent process of sugar fermentation results in accumulation of acids, which dissolve tooth enamel. Especially harmful are suckers and sticky foods like jams, honey, dried fruits and solidified nut pie fillings, which tend to remain on the tooth surface much longer than particle of normal food items. Thus, not only the quantity of sugar consumed but also the exposure of teeth to sugary environment is detrimental for tooth health.

Similarly to alcohol, excessive sugar may interfere with digestive process and displace valuable nutrients. Foods like cakes, candies and colas provide plenty of energy

without supplying noticeable amounts of healthy nutrients. That is they are often called empty-calorie foods. In contrast, foods such as vegetables, fruits and whole grains, which contain large quantities of starches and fibers per unit weight, deliver their glucose together with protein, vitamins and minerals.

A tasty combination of sweet and fatty foods, like ice cream and cheese cakes, is responsible for wide spread consumption of sugar together with fat. Consequently, added-sugar foods are usually high in fat, increasing total energy intake contributing to development of obesity.

It is estimated that in the United States, each person consumes on average about 45 pounds of added sugar annually [15]. This amount translates into about 56 g of added sugar per day is in line with the previous recommendation that refined sugars contribute not more than 10% of total energy intake. Until recently, added sugar consumption at this level has been thought not to pose a major threat to human health with the exception of dental caries. However, the new nutrition guidelines published in January 2005 by the federal government recommended that 10% of total energy intake come from all sugars consumed, which includes both added sugar and sucrose naturally occurring in fruits and vegetables. Therefore, sugar intake has to be reduced whenever it is possible.

1.5.7 Salt

Table salt as a dietary risk factor is mainly associated with sodium intake. However, sodium and chloride in salty solutions have a greater effect on the blood pressure than any other combination of sodium with halogen ions. For years, a high sodium intake has

been known a primary cause of high blood pressure. Similarly, reduction in daily salt intake decreases the risk of hypertension [15].

Sodium is mostly present in processed food such as bread, pies, crackers, chips and pickles. The least amount of sodium per unit weight is found in fresh vegetables, fruits and grains. Current food intake patterns for people in the U.S. are characterized by consumption of about 3100 mg of sodium on a daily basis, which is equivalent to 7.5 g of salt. Studies of control groups indicate that reduction in daily sodium intake from 3450 mg to 1150 mg reduces the systolic blood pressure by 6.7 mm Hg for a typical American diet [20]. Similarly, the same reduction of the sodium intake diminishes the systolic blood pressure by 3 mm Hg for those individuals who are on the balanced diet which emphasizes vegetables, fruits, grains and low-fat products [15]. It should be stressed that there is no salt intake threshold for hypertension. The relationship between the blood pressure and salt intake is direct and progressive [15]. This relationship is in agreement with the trend of salt sales and prevalence of hypertension in the United States. For the last twenty years the amount of food salt sold in this country has steadily increased at the rate of about 4.7% a year. In absolute numbers, table salt consumption almost doubled from 885 tons in 1985 to 1586 tons in 2005 according to American Salt Institute. During the same period, the total number of people diagnosed with hypertension increased from 24% to 29%. Apparently, progress in control in treatment of hypertension could not reverse this trend. Similarly, decrease in salt sales observed in late seventies and early eighties correlated with the downward trend of the total number of high blood pressure patients.

1.5.8 Alcohol

It is estimated that about 10% of men and 3% of women in this country suffer from persistent problems related to alcohol consumption. The autopsy data show that from 10% to 15% of alcoholics has liver cirrhosis at the time of death. Excessive consumption of alcohol (or alcohol abuse) remains a leading cause of morbidity and mortality throughout the world [13].

Although alcohol (also known as ethanol) has a noticeable effect on many vital organs of the human body, its major impact on the human health is associated with the central nervous system and the liver. After being digested alcohol is transported and metabolized in the liver. Alcohol metabolism has a dramatic effect on the liver cell structure. When alcohol enters the liver, the liver cells are forced to metabolize alcohol and their normal function of handling and packaging fatty acids is disrupted. As a result, improperly processed fatty acids may accumulate in large quantities in the liver, setting the first stage of liver deterioration, or the onset of liver cirrhosis – the disease characterized by the liver cell turning orange, hardening and eventually dying.

Narcotic effect of alcohol has been known since pre-historic times. For centuries people used alcohol as anesthetic to relieve pain. It is still widely used during social activities as people believe alcoholic beverages help relax and suppress anxiety. However, it should be noted that brain cells, like liver cells, die when exposed to excessive amount of alcohol.

Currently, it is believed that moderate alcohol consumption, which is about two drinks (one drink equals to 150 g of wine) a day for men and one drink a day for women

doesn't pose a health threat. This assumption is based on the natural ability of the liver to process about 15 g of ethanol per hour without apparent damage to the body. Typically, one glass of wine would contain this amount of alcohol. Unfortunately, observations show that many people who started drinking alcoholic beverages have troubles controlling their alcohol intake. Many people develop physiological dependency on alcohol and need higher and higher alcohol intakes to achieve intoxication. Therefore, they recommend that people attending social gatherings abstain from consuming alcohol whenever it is possible, try to be engaged in an interesting conversation, enjoy eating food, and drink water or non-alcoholic beverages.

1.6 Diet-Related Chronic Diseases

Among the top ten leading causes of death in the United States at least four are believed to be related to diet. These causes include heart disease, cancer, stroke and diabetes and account for about 65%–70% of the nation's two million deaths annually. Fig. 8 shows mortality rates for major chronic diseases. It is seen that the heart disease is responsible for about 285 deaths, while cancer accounts for another 205 deaths per each 100 000 people on the annual basis. Cardiovascular diseases (heart disease and stroke combined) are number one killer and claim 1 death out of every 2.7 deaths. In year 2003 alone, about 1 million people died from heart related problems and stroke in this country. About half out of those deaths was due to coronary heart disease. The good news is that the shift of

American to a healthy diet and advances in medicine help to decrease death rates from cardiovascular diseases by 22% from 1993 to 2003 [12].

When evaluating risk, it is very important to understand that one chronic disease often serves as a risk factor for another and vice versa. As an example, Fig. 9 shows complex interrelationship between several leading chronic conditions [13] with arrows pointing from risk factors to diseases. Thus, a person with atherosclerosis and hypertension has a high risk of suffering from stroke. On the other hand, a diabetes patient is more likely to develop heart disease. Similarly, an obese person has a high chance of developing atherosclerosis and hypertension and so forth. Nonetheless, primary culprits remain to be behavioral risk factors which are mostly diet-related. The sections provide a brief overview of chronic diseases with the emphasis on their relationship with diet.

1.6.1 Coronary Heart Disease

Coronary heart disease (CHD) is one of the major members of the group of illnesses, which are collectively called cardiovascular diseases (CVD). The heart disease affects about 13 millions of Americans [12] and represents the leading single cause of death in the world. In particular, the heart disease contributed about 27% to the total mortality rate in the US in years 2001-2003. An average, every minute someone dies from coronary events in this country. The symptoms of the heart disease include myocardial infarction or heart attack and angina pectoris, also known as chest pain.

Among factors contributing to increasing the risk of coronary heart disease are smoking, physical inactivity, obesity, and high blood pressure. The risk increases with age and statistically is higher for male population and some ethnic groups.

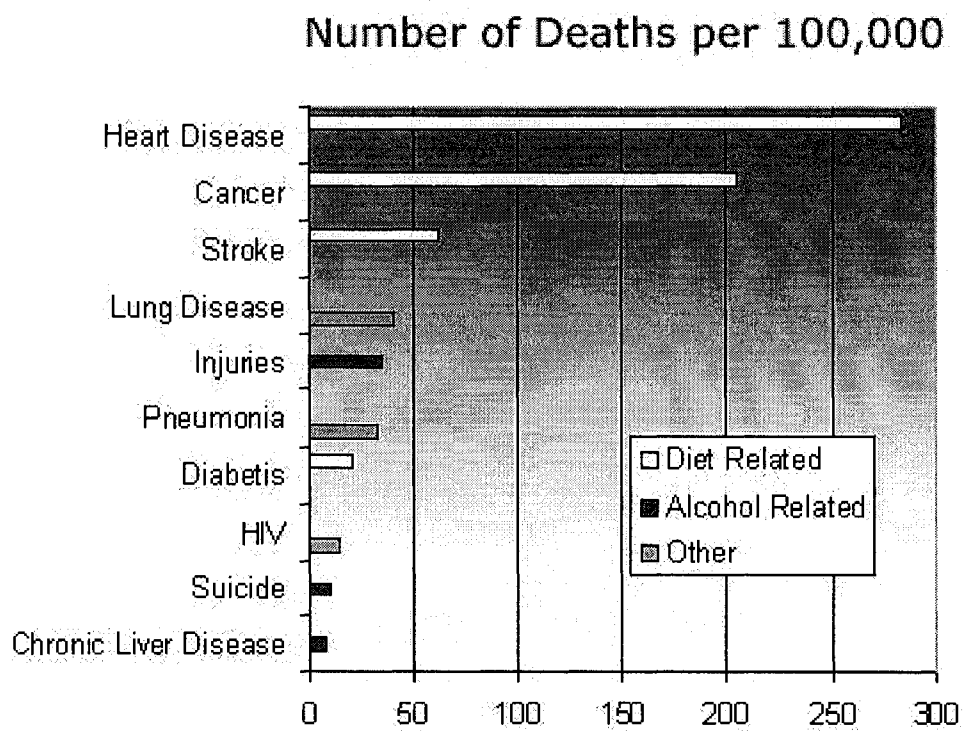


Figure 8: Leading Causes of Death in the United States [13].

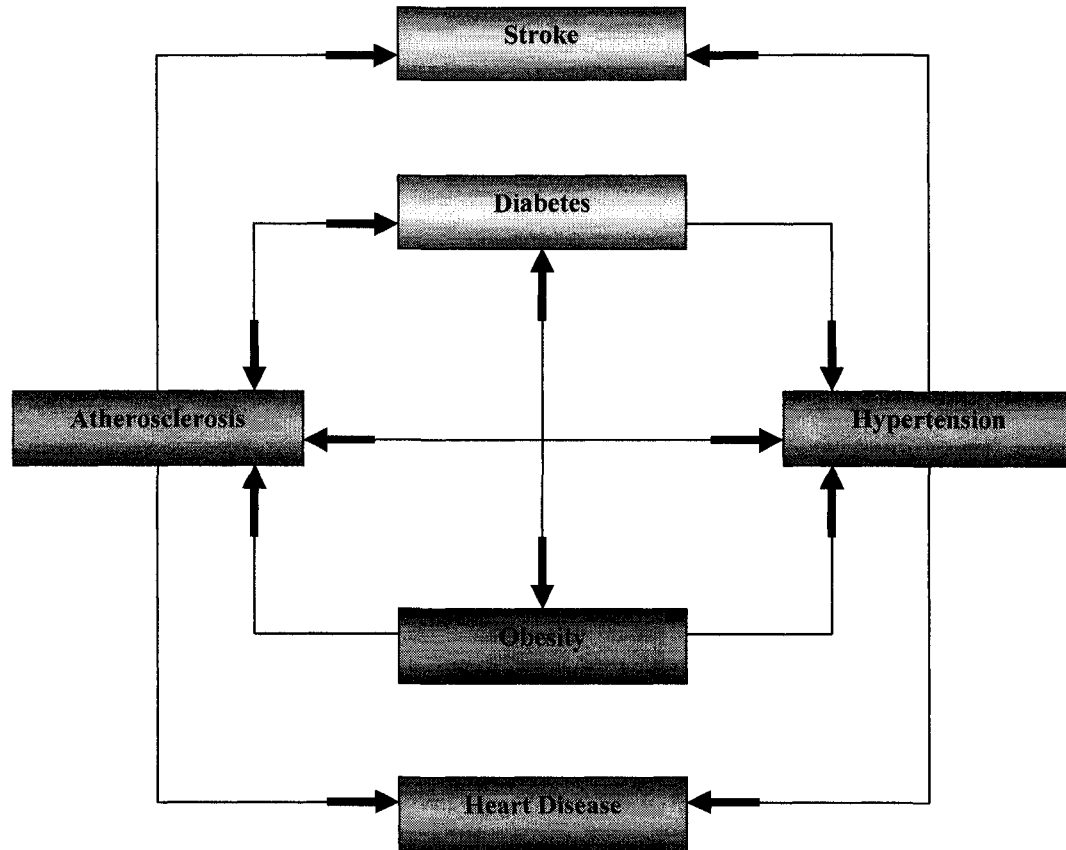


Figure 9: Interdependence of Major Chronic Diseases and Risk Factors.

As it follows from Fig. 9, contracting a certain disease may also increase the risk of coronary heart disease and vice versa.

Physiologically, the heart disease involves development of atherosclerosis and hypertension, which in turn exacerbate each other. The initial stage of atherosclerosis is associated with formation of soft fatty streaks - plaques on the inner walls of arteries and especially at the artery branch points. Although plaque formation may occur in any blood vessels, it mostly affects the coronary arteries. The coronary arteries are the blood vessels that supply oxygen and nutrients to the heart muscle tissue – myocardium. The

myocardium is so thick that it requires its own vessel system for blood delivery. That is why the heart disease is often called coronary heart disease. In course of time, the plaques tend to absorb minerals, become larger and harden. As a result, the blood vessels become narrower, restricting blood flow to the heart muscles and limiting supply of oxygen to the cells. In addition to this effect, plaques on the blood vessel walls attract small, cell-like bodies in the blood, which are called platelets. The natural function of platelets in the body is to form clots around the injured areas. Usually clots form and dissolve at the same rate so that there is a safe balance in clot formation. However, the presence of plaques on the walls of arteries trigger a faster formation of clots since the body recognizes the plaques as injuries. Abnormal clot generation may result in clot sticking to a plaque in a blood vessel. The clot can then grow larger to restrict or even close off the vessel. In another scenario the clot may be blown away from its initial position and travel through the circulatory system to land in a vessel of a smaller diameter where the clot cannot go through. In both cases a life-threatening event – thrombosis develops, which may result in completely shutting off supply of oxygen and nutrients to the adjusting tissues. Another adverse effect of plaques is increasing the blood pressure. Restriction of the blood flow by plaques and clots forces the heart to generate more pressure to maintain adequate supply of nutrients by the blood to the tissues. The higher blood pressure in turn increases the probability of plaque formation since the initial plaque constituents bombard the vessel walls at a higher frequency. Thus, development of atherosclerosis is self-sustaining and self-accelerating process which makes it even harder to cure.

According to [21, 22], “nine easily measured and potential modifiable risk factors account for over 90%” of the total risk of suffering from myocardial infarction. Those risk factors are consistent across different geographic regions and ethnic groups and are almost independent of gender. They include (1) cigarette smoking, (2) elevated blood lipid levels, (3) high blood pressure, (4) diabetes, (5) abdominal obesity, (6) physical inactivity, (7) low daily intake of fruits and vegetables, (8) over-consumption of alcohol, (9) psychosocial index. The latter is based of ratings of stress, well-being, psychological distress, and illness behavior. From this it follows, that diet high in saturated fat and cholesterol and low in vegetables and fruits is identified as a major risk factor and contributor to progression of atherosclerosis. Other diet-related risk factors are hypertension, diabetes and excessive alcohol consumption. In 2006 direct and indirect cost of coronary heart disease amounted to \$142.5 billion and is the highest among chronic diseases.

1.6.2 Stroke

Stroke is another representative of the cardiovascular diseases and is responsible for about 6.6% of all deaths in the United States. If considered separately from other cardiovascular diseases, stroke would be ranked the third leading cause of death after heart disease and cancer. On average, a stroke occurs every 45 seconds. Every 3 minutes someone dies from stroke in this country with women being disproportionately affected as compared to men. Of every 5 deaths from stroke, 3 occur in women and only 2 in men.

Stroke is associated with damage of the arteries leading to and within the brain. A stroke occurs when a blood vessel that carries oxygen and nutrients to the brain is either

blocked by a clot or bursts. When that happens, a part of the brain cannot get the blood (and oxygen) it needs and starts dying since even few-minute interruption in supply can cause damage to nerve cells in the brain. The disease can be caused either by a clot obstructing the flow of blood or by a blood vessel rupturing and preventing blood flow to the brain [12]. They distinguish several kinds of stroke. In the so-called thrombotic stroke, which accounts for 60% of cases, a blood clot occurs in the artery feeding the brain, causing brain starvation or ischaemia (transient ischaemic attack). The embolic stroke is similar to the thrombotic one except for the clot being originally formed elsewhere in the body and then transferred by the blood stream to the brain, where it eventually blocks a blood vessel. This type of stroke accounts for 20% of all stroke conditions. If a small artery penetrating the brain tissue is blocked by a clot, a lacunar stroke can happen, causing damage in tiny areas adjacent to the affected artery. The lacunar strokes occur in 5-10% of cases. Finally, the remaining 10-15% of stroke cases is related to two types of brain bleeding. One is subarachnoid haemorrhage, or bleeding into space between the brain surface and the skull due to busting of a weakened blood vessel. Another is intracerebral haemorrhage associated with bleeding deep inside the brain. Being a disease of the blood vessels and the brain, stroke effect on the patient is tremendously different from that associated with the heart disease.

The stroke condition can affect the entire human being with a common disability, which manifests in complete paralysis on one side of the body. A related disability that is slightly less debilitating than paralysis is one-sided weakness. Stroke may frequently cause problems with thinking, awareness, attention, learning, judgment, speech

understanding and memory. Another consequence of stroke is emotional, as many stroke survivors suffer from depression. Stroke patients often have difficulty controlling their emotions or may express even inappropriate emotions and change in character. People, which have suffered a stroke, may have numbness or strange sensations. Generally, the stroke risk factors are similar to those for the heart disease. Having a heart disease, specifically hypertension, is considered an additional stroke risk factor, which is more probable with an advanced age. That is why the probability of stroke increases with age almost exponentially with 90% of all stroke cases reported in people who are 55 years of age and older. Interestingly, men are more likely to suffer from stroke than women. However, since women on average live longer, the period of time they may experience a stroke is larger than that of men. As a result, women outnumber men in total number of strokes by the ratio 3:2.

1.6.3 Hypertension

Hypertension, or elevated blood pressure, is one of the most common cardiovascular diseases among Americans. After age 55, more than 46% of men and 55% of women have an elevated blood pressure. High blood pressure was primary or contributing cause of death of 11% of those who died in 2003. Hypertension is defined as systolic pressure being higher than 140 mm Hg, or diastolic pressure of 90 mm Hg or higher. About 28% of American adults age 18 and older, or 59 million people, have pre-hypertension, a condition characterized by a blood pressure slightly higher than normal and being a precursor to hypertension. Thus, hypertension affects a significant percentage of adult population.

Although, hypertension doesn't pose an immediate death threat to a patient, is closely related and considered as a major risk to many life-threatening conditions such as heart attack, heart failure and stroke. Elevated blood pressure may also result in injuries of eyes and kidneys. However, the most important implication of hypertension is the way it affects the work of the heart muscle – myocardium. When the blood pressure is higher than normal, the heart has work against an additional force to maintain an adequate supply of blood to the vessels. This additional effort requires extra muscle energy. As a result, the heart muscle tends to thicken, leading to the so-called heart hypertrophy. Since the bottom chambers of the heart are responsible for pumping blood into lungs and body, the lower part of the heart becomes larger. Unfortunately, the heart response to the increased blood pressure by enhancing the myocardium often leads to overall weakening of the heart muscle due to failure to meet the increased demand for oxygen and nutrients as well as problems of synchronization of the atria (upper chambers) and the augmented ventricles (lower chambers). Weakening of the heart muscle, in turn, causes a smaller supply of blood per unit time to the essential parts of the body including the heart itself and the brain. That is why hypertension has exacerbating effect on development of coronary heart disease and stroke. The estimated cost of hypertension in the United States in 2006 is about \$63.5 billion and exceeds that of stroke (\$57.9 billion). The risk factors of hypertension include all of those for the heart disease. In addition to that, table salt intake has a well-pronounced and an aggravating effect on hypertension [15].

1.6.4 Obesity

Obesity is considered as epidemic in the United States. Its prevalence has increased dramatically since 1991 and continues to be on the rise. Physiologically, obesity is defined as excessive accumulation of body. Typically, 20% or more over an individual's ideal weight is considered to be obesity. Current guidelines for obesity are based on the body mass index (BMI). The BMI has a dimension of kg/m^2 and is equal to the body weight divided by the height squared. Obesity corresponds to BMI values of 30 or higher. People with the body mass index between 25 and 29 are considered overweight. It is estimated that full 50% of all Americans are either overweight or obese. The percentage of obese people in the age category of 15 years and older is about 30%.

Despite the ongoing debate on whether to include obesity into the group of chronic diseases or not, the World Health Organization named obesity worldwide epidemic. Impact of obesity on the society cannot be overestimated. Indeed, obese people are more likely to experience illness symptoms like fatigue and wheezing and suffer from cardiovascular diseases, cancer, and diabetes mellitus.

Causes of obesity can be split into three categories – genetics, overeating, and physical inactivity. Obese gene inheritance accounts for about 10% of obese adult people. Overeating and physical inactivity are responsible for the rest of the obesity cases, which is about 20% of all population aged 15 and older. Among other risk factors, researchers have found that there is a positive correlation between fat intake and body mass index. On the contrary, a diet high in carbohydrates and fiber negatively correlates with body fat. Strong evidence also suggests that children who eat high fat food are more likely to

be overweight and obese [23, 24]. Thus, reducing fat in the diet plays a significant role in weight loss and obesity treatment.

1.6.5 Diabetes Mellitus

Diabetes mellitus is a chronic condition characterized by elevated blood glucose and either insufficient or inefficient insulin, which is a hormone responsible for control of glucose transport from bloodstream to cells. Diabetes moved from the seventh leading cause of death in the United States in year 2000 to the fifth in 2005. In addition, it underlies or contributes to a variety of other diseases, including heart disease, stroke and hypertension. People with diabetes are twice as likely to develop these cardiovascular problems as those without diabetes [15]. Each year about 85 000 people in this country undergo non-traumatic lower-limb amputations resulting from diabetes complications. According to American Diabetes Association, there are 20.8 million children and adults in the United States, or 7% of the population, who have diabetes. It is believed that 6.2 million people out of 20.8 with diabetes are unaware of their condition [25]. Overall, diabetes increases the death rate among people by a factor of two as compared to that of people without this disease.

There are two types of diabetes. In type I diabetes, the pancreas are unable to synthesize the hormone insulin. As a result, the energy metabolism in a human body is severely distorted. To survive under this condition an individual needs constant injections of insulin. Type I diabetes is relatively rare and accounts for 5 to 10% of all cases.

The predominant form of the disease is type II diabetes, which develops in people aged 40 and older. In type II diabetes, pancreas can produce insulin, but its effect on the

glucose level is often inadequate due to reduced cell sensitivity to insulin. As a result, the cells don't absorb glucose from the blood quickly enough and the glucose level in the blood continues to rise. This stimulates pancreas to make even more insulin. Eventually, the ability of pancreas to produce insulin may be exhausted, leading to a self-aggravating condition.

When cells don't get sufficient amount of glucose from the blood, they notify the brain about being hungry. A person feels excessive appetite and eats more food than actually necessary (polyphagia), which frequently results in weight gain. However, over the long term most important ramifications of diabetes are related to development of chronic conditions, triggered by unregulated glucose levels in the blood. Systematically elevated glucose concentration damages the vessel walls and nerve structure, resulting in poor circulation, impaired sensation and pain in hands and feet. Diabetes combined with atherosclerosis causes a gradual reduction in the blood supply to the feet tissues, leading to degenerative changes in limbs and gangrene. People with diabetes contribute to 60% of all lower-limb amputations performed annually. Another complication of diabetes is impaired vision and blindness (diabetic retinopathy) among the adults aged 20 year and older, which accounts for about 20 000 new cases every year.

Prevention and control of diabetes relies to a great extent on dietary measures. It is found that people with both types of diabetes can reduce micro-vascular complications (affecting eyes, kidneys and nerve structures) by 40% when they carefully control intake of foods with high triglycemic index. Following a low-fat diet helps to better monitor blood lipids (HDL and LDL) and reduces cardiovascular complications by 20% to 50%

for diabetes patients. Similarly, hypertension control through diet and medications can reduce the risk of macro- and micro-vascular complications in persons with diabetes by 30% to 50% [26].

1.6.6 Osteoporosis

Osteoporosis is a disease characterized by a dangerously low bone density and deterioration of the bone tissue. The condition results in bone fragility and increased risk of fracture. In this country, osteoporosis causes 1.5 million fractures every year, including 300 thousand of devastating hip fractures.

In October 2004, U.S. Surgeon General published a report with a warning that by 2020, half of all Americans older than 50 will be at risk for fractures from osteoporosis and low bone mass if no immediate action is taken by individuals at risk, doctors, health systems, and policymakers [27]. It is estimated that about 10 million of Americans aged 50 and older have osteoporosis and another 34 million have a reduced bone density and are at risk of developing the disease.

Few people realize how dangerous the consequences of osteoporosis are. Meanwhile, statistical data indicate that 20% of patients suffered from hip fractures die within the first year after the fracture occurred due to complications related either to the fracture or surgery to repair it. "I always worried about heart disease and cancer, but was never concerned about the health of my bones," said Abby Perelman, who is being treated for osteoporosis, and added "I wish I knew then what I know now -- that a healthy diet and physical activity can make bones stronger and healthier" [27].

The cost to the healthcare system associated with osteoporosis fractures in the U.S. is estimated to be about \$20 billion on the annual basis and is expected to rise dramatically by the year 2030 [28].

It is recognized that, although bone loss cannot be stopped, it is never late to protect bone health by ensuring adequate intake of calcium and vitamin D, as well as performing weight-bearing exercises [28]. That is why healthy intakes of calcium and vitamin D are considered to be at the base of a pyramid for maintaining bone health and preventing fractures.

1.6.7 Dental Caries

Dental caries, or tooth decay, is characterized by formation of dental cavities or holes which damage the tooth structure. The disease is widespread across the globe and may lead to acute pain, inflammation, tooth loss, and even death. Estimated nine out of ten schoolchildren and overwhelming majority of adults have experienced dental caries. The disease is especially common in Asian and Latin American countries. In the United States, dental caries is the first chronic disease children encounter in their lives and the most important cause of tooth loss in younger people [13].

The major cause of caries is accumulation of plaque on the tooth surface. The plaque is a sticky substance representing a combination of food remnants, bacteria, which is always present in the mouth, acids produced by bacteria as a result of food fermentation and saliva. Plaque begins to accumulate on teeth within 30 minutes after food intake. If it is removed from the teeth during this period, tooth decay can be significantly reduced. The plaque, which remains on the tooth surface, mineralizes into

tartar and contributes to the development of caries. It also irritates the gums, resulting in gingivitis and later in periodontitis. Chemically active plaque acids dissolve the tooth enamel, creating holes in the tooth tissue (cavities). Normally, cavities don't cause pain until they grow large enough inside the tooth to destroy the tooth nerve and blood vessels. Untreated tooth decay also destroys the internal structures of the tooth (pulp) and eventually results in the loss of the tooth.

Early in the twentieth century, American practitioners noticed that water fluoridation reduced the risk of dental caries. Since then water fluoridation along with fluoride tooth paste has been effectively used as caries prevention means. Although, fluoridation may decrease the prevalence of dental caries by 50% among children and 30% among adults in some cases, it cannot fully eliminate occurrence of the disease. Refined carbohydrates (industrially processed sugars and starches), sticky foods and frequent snacking increase the risk of tooth decay, since all of those factors facilitate plaque generation by bacteria. In short, reducing sugar and sweets intakes and carefully monitoring overall frequency of meals and snacks can help significantly decrease the probability of dental caries development.

Diverticular disease comprises two closely-related conditions diverticulosis and diverticulitis. Diverticulosis is a condition in which small pouches or pocket-like openings develop in weakened areas of the intestine. The name of the disease originates from the Latin word diverticulum or small pouch. A person with diverticulosis suffers from the stool being too hard and bowel muscles having to strain to move it along.

Excessive muscles straining increases pressure in the intestine (mostly colon) and leads to formation of pouches in weak spots along its walls [13].

About 10 percent of Americans over the age of 40 have diverticulosis. The condition becomes even more prevalent as people age. About half of all people over the age of 60 in this country have diverticulosis. Diverticula may become infected or inflamed, causing the condition called diverticulitis. This happens in 10 to 25 percent of people with diverticulosis [29].

Diverticulosis) is most common in developed or industrialized countries. The disease is relatively rare in countries of Asia and Africa, where people eat high-fiber vegetable diets. Diverticulosis has become a problem in the United States since the early 1900s with the advent of processed foods, which were quickly integrated with the traditional American diet. At that time, intake of popular high-energy and low-fiber food items like refined sugars and starches increased at the expense of vegetables and fruits. Simultaneously, whole grain cereals were replaced by white flour. As a result, the amount of fiber in the typical American diet decreased significantly.

Initially, most people who have diverticulosis do not suffer from any apparent symptoms. Still symptoms such as bloating and constipation can be early signs of diverticulosis development. As diverticulosis progresses the probability of diverticulum inflammation increases. A sudden abdominal pain may be an indication of the onset of diverticulitis caused by infectious bacteria, which dwells in the small pouches [29]. Other symptoms of diverticulitis include fever, nausea, vomiting, chills, cramping, and constipation. The severity of symptoms depends on the extent of the infection and

complications. In some cases, an attack of diverticulitis is serious enough to require a hospital stay and possibly surgery.

Increasing the amount of fiber in the diet may reduce symptoms of diverticulosis and prevent complications of diverticulitis. Sufficient amount of fiber keeps the stool soft and lowers pressure inside the colon so that bowel contents can move through easily [30, 31].

1.7 Overview of Existing Food Pattern Calculators

Despite noticeable public interest in healthy lifestyle and recommendations on reducing disease risk, existing software tools often fall far short of user expectations. There are two major groups of software tools used for food intake assessment. One is concerned mainly with body weight reduction and calorie intake calculations. The other group of tools provides the possibility to select a variety of meal ingredients and foods servings. After the choice is made, the user is provided with calculated amounts of essential nutrients in the selection. Commercial products and web sites with food intake calculators try to combine both features, revolving mainly around the need to loose weight. Below a brief overview is made of some of the tools available on the Web and as standalone applications.

A simple nutrition calculator (“Nutritional Value Calculator”, November 2006, <http://www.3.waisays.com/calculator.html>) allows one to select up to 16 food items and then calculates the cumulative amounts of different nutrients versus Recommended Daily

Allowances (RDA). The listing of nutrients is pretty extensive and includes about 150 nutrients. However, for almost 80% of these food composition items is not available. That leaves the user wondering whether the data is not known or simply unavailable in the tool database.

The web site “The Sonoma Diet” (The Sonoma Diet, November 6 2006, <https://www.sonomadiet.com/publicsite/funnel/register.aspx>) offers not only healthy ways to lose weight, flavorful and easy-to-make recipes, food intake planners, but also shopping lists. Unfortunately, the site doesn’t provide the possibility of the free trial tool evaluation.

Another web site - Myfooddiary.com (“A New & Effective Approach to Weight Loss”, November 6, 2006 <http://www.myfooddiary.com>) provides an easy-to-use on line food diary and a number of useful information resources. The user can search the food database of over 40,000 items and add his/her items to the food diary. The tool generates recommendations for the user on how to improve the diet. Apart from calculating total calories, the system tracks the intake of fat, saturated fat, carbohydrates, sugars, fiber, sodium, cholesterol, vitamin A, vitamin C, iron, calcium, and water.

The government-sponsored Center for Nutrition Policy and Promotion page (“Center for Nutrition Policy and Promotion”, November 7, 2006, <http://www.mypyramidtracker.gov/hei.aspx>) leads the user directly to the food item search utility. The search engine, however, may not be smart enough to discern “bread” and “breaded chicken breast”. Therefore, search for bread results in the listing that contain bread, as well as breaded meat and poultry dishes. The food analyzer provides

generic guidelines for 5 food groups (grains, vegetables, fruits, milk and meat/beans) and few nutrients such as different types of fat, and sodium and risk-related ingredient such as cholesterol.

DietPower.com (DietPower – Personal Weight and Nutrition Software”, March 3, 2008, <http://www.dietpower.com>) offers an interesting standalone program “Diet Power”, which can be downloaded for 15-day free trial. The program works independently or in conjunction with different popular diets such as the low-carb, Atkins, American Heart Association diet and others. The tool tracks intake of 33 nutrients, which is significantly more than typical food intake calculators do, and contains a database of about 11,000 food items. The program is able to learn favorite intake patterns of the user. It coaches the user and makes suggestions on balancing nutrient intake to reduce the risk of chronic diseases. Apart from registering daily food intake patterns, the DietPower main menu contains options for moving personal records to a backup drive or email, working with recipes, recording physical activities, looking up for food items, and other useful buttons, including a link to DietPower web site updated every week. Especially helpful is the diary menu, which allows one to go back and see the food intake patterns for the previous days. Searching for food items is relatively fast and the search can be narrowed within a certain category automatically without user intervention. The tool allows the user to specify a favorite set of items, which is transferred from one day to another, so that there is no need to select them again. As opposed to vast majority of other tools, Diet Power provides the possibility to drag and drop a food item from the lookup selection list to the food intake pattern pane for a specific day, which is pretty handy and creates a positive

user experience. Right-clicking on the selected food item opens a separate window with the item's nutrient information, so that the user can quickly decide, based on the nutrient data, whether he/she is interested in this particular item, before dropping it into the daily intake pattern. Another useful feature is the Nutrient Quotient (NQ) which displays the amount of different nutrients in a daily intake pattern in a form of a table or chart.

Summarizing this brief overview, it should be stressed that none of the existing tools provides a possibility for quantitative direct or reverse health risk assessment related contracting a chronic disease due to diet risk factors. Most of the tools require a lot of time and dedication from the user to achieve meaningful results. The user often prompted to go through numerous and sometimes tedious options and questions before he/she gets the information of interest. The food intake analysis, as a rule, is reduced to a simple listing of tens of essential nutrients whose meaning and immediate impact are not always clear to the end-user.

Meanwhile recent advances in medical statistics and nutrition science make it possible to quantitatively relate dietary habits and risk of developing a chronic disease. Especially interesting is the idea of computing various food intake scenarios and linking them to the risk of contracting a cardiovascular disease. This idea has led to the design and implementation of a reverse risk assessment system, which are described in Chapter 2 and Chapter 3, respectively. The system implemented as a web-based application which allows the user to establish a direct connection between an accepted level of risk for developing a cardiovascular disease (heart disease, stroke, and hypertension) and diet-related risk factors.

2. System Design

This chapter describes the application design and the mathematical model used for risk assessment calculations.

2.1 System Requirements Overview

The risk assessment software is conceived as a versatile tool for making quick and efficient estimates of cardiovascular disease risk associated with a certain diet. The tool is also supposed to serve as an informational resource on diet-related risk factors and chronic diseases.

The requirements for the application are listed in Table 1 and include the capabilities of specifying disease risk, intake of energy, fat and sodium. It is also assumed that the user is able to account for three behavioral risk factors such as smoking, obesity, and physical inactivity when estimating the probability of contracting the disease and retrieving a respective set of food intake patterns. In the majority of cases, the system should be able to generate a statistically significant number of food intake patterns for a particular set of search criteria.

Table 1: System Requirements

No.	Description	Priority
Risk Assessment Capabilities		
1	The application must be able to generate a set of food intake patterns associated with a certain disease and its risk.	High
2	The user must be able to specify a degree of acceptable risk associated with one or several diet-related chronic diseases.	High
3	The user must be able to specify a particular cardiovascular disease associated with the specified risk.	High
4	The user must be able to specify calorie intake in the food intake patterns.	Medium
5	The user must be able to specify fat and saturated fat intake in the food intake patterns.	High
6	The user must be able to specify sodium intake in the food intake patterns associated with the certain degree of disease risk.	High
7	The user must be able to specify a diet type (Traditional American, Vegetarian, Mediterranean or Vegan).	Medium
8	The user must be able to account for behavioral risk factors like smoking, obesity and physical inactivity.	High
9	The user must be able to specify the number of cigarettes per day when accounting for the effect of smoking on the disease risk.	Medium
10	The user must be able to specify the Body Mass Index (BMI)	Medium

	when accounting for the effect of obesity on the disease risk.	
11	The user must be able to specify exercise time per week when accounting for the effect of physical inactivity.	Medium
12	The application must provide the search capability to select those food intake patterns which meet the specified search criteria	High
13	The system must be able to generate a statistically significant number of food intake patterns for a particular set of search criteria.	Medium
Food Intake Patterns		
14	Food intake pattern description must provide an itemized listing of food items/products and associated food groups	High
15	Food intake pattern description must provide details about intake of vitamins for specific food items, as well as the whole pattern.	High
16	Food intake pattern description must provide details about intake of minerals for specific food items, as well as the whole pattern.	High
17	A food intake pattern must provide details about intake of fat/saturated fat for both food items and the whole pattern.	High
18	A food intake pattern must provide details about intake of carbohydrates for both food items and the whole pattern.	High
19	A food intake pattern must provide details about intake of fiber for both food items and the whole pattern.	High
20	Food intake pattern page must provide convenient navigation	Medium

	across informational data which characterizes a pattern.	
21	The food pattern search results must be presented in a table which displays food intake summary and basic food intake properties.	Medium
22	The food intake table must allow the user to sort results by total food, calorie, fat/saturated fat, and sodium intake.	Medium
	Informational Support	
23	The application must provide links to external web sites with relevant nutrition and healthy lifestyle information.	Medium
24	The application must provide a glossary of major terms and abbreviations related to chronic diseases, diet and nutrition.	Medium
25	The glossary should provide easy navigation across different terms ordered alphabetically.	Medium
26	The application should provide basic information about major diet-related risk factors.	Medium
27	The application should provide basic information about major diet-related chronic diseases.	Medium
28	The application must provide information about incoming conferences and events related to nutrition and health care issues.	Low
	Application Access and Security	
29	The access to the application must be protected by user ID and password.	Medium
30	The user must be able to update his/her profile	Low

31	The system administrator must be able to create a new user profile	High
Non-Functional Requirements		
32	The response time for pattern search must be reasonable and should not exceed 3-4 seconds.	High

2.2 Use Case Model

The Use Case Model logically follows from the application requirements specified in the previous section and includes use cases for (a) performing basic risk assessment operations (like selecting a chronic disease, specifying the probability to contract the disease, and setting diet preferences and behavioral risk factors), (b) browsing and exploring food intake patterns, (c) reading information about diet risk factors and chronic diseases, and using glossary and external links. In addition to that, there are several use cases related to accessing the application, creating new and updating existing user profiles.

Figure 10 illustrates the use case scenarios for a standard user, while Figure 11 shows use cases for a system administrator. The key use case “Calculate Risk” is conveniently split into two use cases which pertain to selecting risk criteria and searching for patterns. A detailed description of uses cases is provided by a set of use case sequence diagrams which are presented in Appendix A.

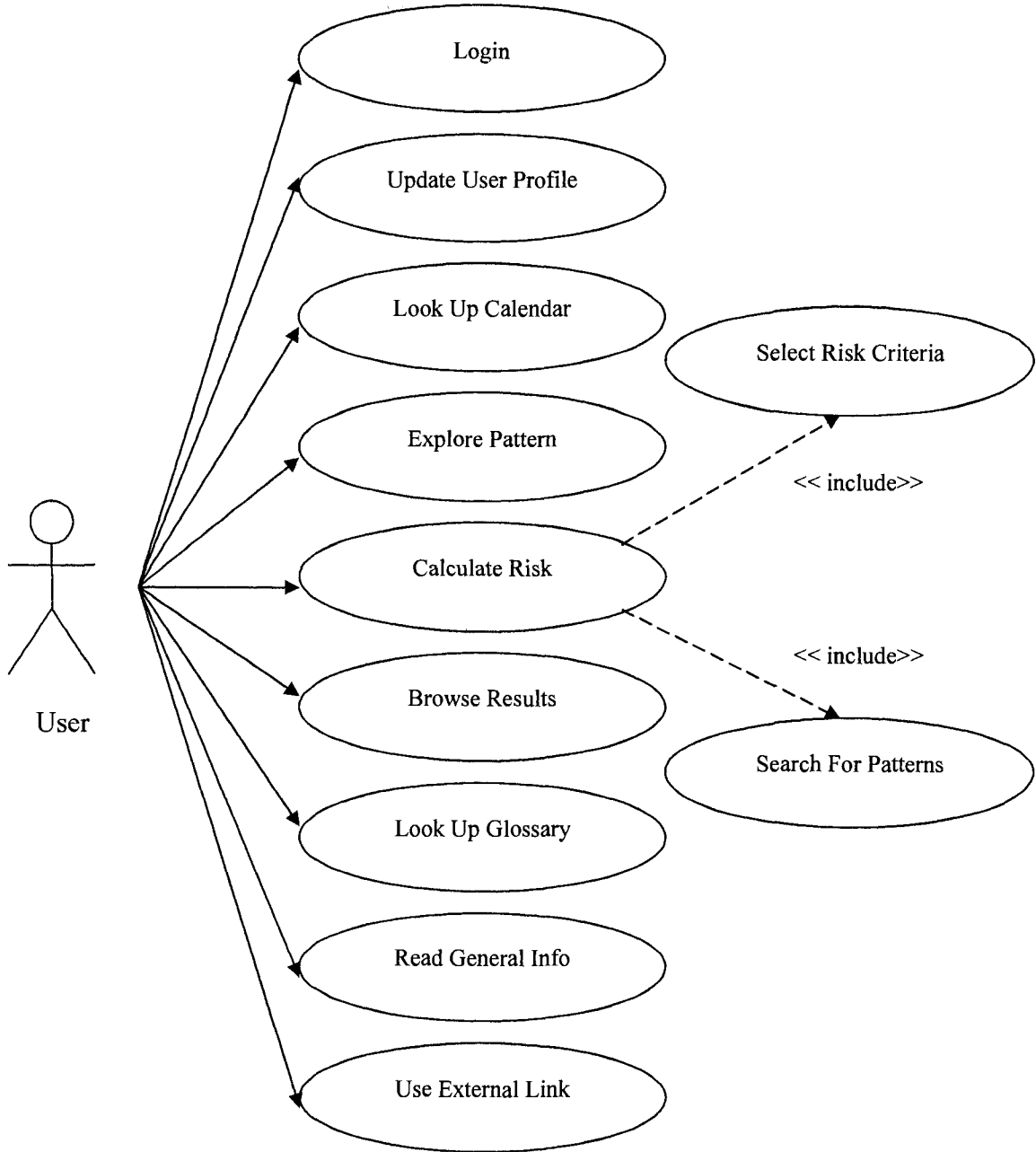


Figure 10: Basic Risk Assessment Use Case Diagram

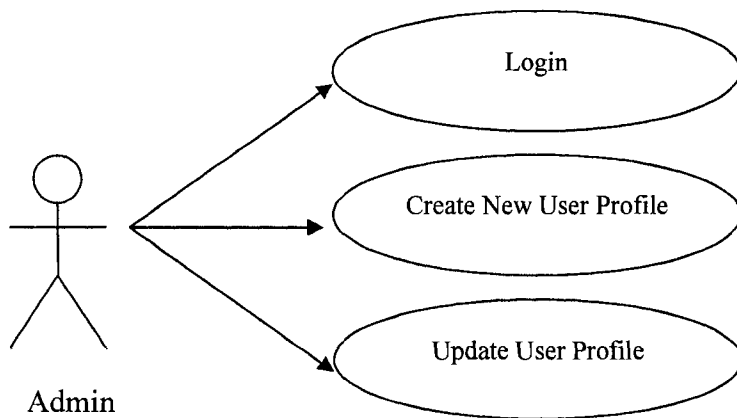


Figure 11: System Administrator Use Case Diagram.

As it follows from the system requirements and the use case model, the key feature of the application is the ability to generate a set of food intake patterns which satisfy the disease risk criteria. However, those risk criteria themselves cannot be used directly for food intake pattern selection. Instead, they need to be interpreted and linked to the probability of contracting the selected disease. The relationship between the risk and diseases is provided by the mathematical model discussed in the next section.

2.3 Mathematical Model

This section is devoted to development of the mathematical model for calculating a probability to contract a chronic disease in a lifetime. Due to complexity of the chronic disease risk assessment problem, we consider first non-diet risk factors and then address

the issue of diet-related risk factors. The end result is an expression for the probability of contracting a chronic disease, which accounts for both types of causal risk factors.

Although some results have a general nature, our major focus is on cardiovascular disease risk assessment. One of the reasons is that there is enough statistical data to test and justify assumptions underlying the model. Most importantly though, cardiovascular diseases are the major killer in the developed countries, as there is growing scientific evidence that rise in prevalence of CVD reflects “...a significant change in diet habits, physical activity levels, and tobacco consumption worldwide as a result of industrialization, urbanization, economic development and food market globalization” [32].

2.3.1 Relative Risk of Non-Diet Risk Factors

The relative risk is defined as the ratio of the incidence in the exposed population over the incidence in unexposed persons [33]. Although the notion of relative risk is convenient and seems to be intuitively clear, this definition is not entirely precise. In particular, the denominator of the ratio may be either the average risk of the entire population or the risk of a group devoid of risk factors [33]. It is pretty obvious that a population group completely devoid of risk factors would exhibit a zero or, otherwise, very small probability to contract a disease, resulting in a disproportionately high value of the relative risk. Therefore, the definition of the relative risk requires further elaboration. We assume that denominator of the ratio is the average risk of a very large group of people who are devoid of non-diet or behavioral risk factors, but in all other aspects are similar to the population group exposed to a particular risk. For instance, we assume that

both exposed and non-exposed groups follow the same traditional diet. Thus, the general expression for the relative risk reads:

$$(1) \quad R_i = p^{-1} (p + q_i)$$

where p is the incidence of disease in general population devoid of behavioral risk factors and q_i is the increase probability due to the risk factor i .

The continuous process of food digestion takes place against the background of simultaneous exposure to non-diet related risk factors. It is logical to assume that non-diet related risk factors like smoking, being overweight and others influence the body metabolism by increasing relative risk of contracting a chronic disease. Medical studies conducted since the beginning of the twentieth century confirm the link between diet and behavioral risks. The report published by the World Health Organization emphasizes this idea, stressing that "...risk behaviors, such as tobacco use and physical inactivity, modify the result" of food digestion "for better or worse" [34]. Below, the analysis is focused on three behavioral risk factors such as smoking, physical inactivity and obesity, and two non-modifiable risk factors like gender and ethnicity.

The detrimental effect of smoking on human health was observed as early as 1908 [35]. Since then it has been found that the relative risk of cardiovascular disease among smokers is 2-4 times greater than in non-smokers [36, 37]. The relative risk increases from 2 to 4 depending on the number of cigarettes smoked a day. In particular, the risk is the largest for heavy smokers (more than 40 cigarettes a day) and smallest for light smokers (10 or fewer cigarettes a day) [37]. Employing a linear approximation, the relative risk of smoking can be described as

$$(2) R_s = 0.075N + 1$$

where N is the number of cigarettes smoked per day.

Physical inactivity is another risk factor, which has a comparable relative risk. According to the World Health Organization [34] people who are physically inactive have elevated risk for cardiovascular diseases, like heart disease and stroke. The average relative risk factor for physical inactivity is estimated to be 1.5, although in some countries it is believed to be about 2. People are considered to be physically inactive if they spend less than 2.5 hours per week on physical activity, which is at least 2.5 hours of vigorous physical exercises. Negative effects of physical inactivity include lower metabolism and weight gain, less efficient blood circulation and increase in blood clot formation in vessels. At present, only 32% of white people and about 25% of African Americans in this country meet the recommended level of physical activity per week. Taking into consideration the definition of the physical inactivity and statistical evidence on the respective relative CVD risk factor, the dependence of the relative risk on the hours of physical activity can be approximated by the formula

$$(3) R_p = -T/150 + 2.0 \quad (T \leq 150),$$

where T is the number of minutes spent on physical activity per week during leisure time. As it follows from formula (2), the relative risk reaches maximum value of 2 for entirely inactive people and approaches unity at $T \sim 150$, which corresponds to about 2.5 hours of physical exercises per week. It is also assumed that further increase in time spent on physical exercises has little impact on the risk factor R_p , which is about unity ($R_p = 1$) for

T > 150 minutes per week. Sedentary lifestyle often leads to obesity, which is another serious risk factor

Obesity influences the relative risk of cardiovascular disease slightly differently for men and women. The probability to develop the disease is by 20% higher for obese women than for obese men [38, 39]. On average, the probability increases almost linearly from overweight to extremely obese people with the relative risk reaching the value of 2.2 [40]. Thus, it is reasonable to approximate the obesity relative risk by a linear dependence starting at 1 for normal weight (BMI = 18.5-24.9) and going up to 2.2 for extreme obesity (BMI > 40). Equation (3) describes the dependence of the obesity relative risk as a function of body mass index (BMI):

$$(4) R_o = 0.094B - 1.34 \quad (B \geq 24.9)$$

where B is the body mass index. When deriving expression (3), it is assumed that the body mass index in the range from 18.5 to 24.9 doesn't affect the risk of cardiovascular disease and, therefore, the relative risk factor for this range is equal to unity. On the other hand, in the case of extreme obesity (BMI ~ 40), the relative risk may become greater than 2.0. Note that, for the average value of BMI = 29.2, formula (4) yields the relative risk $R_o = 1.4$, which is good agreement with data available in literature [41-44].

Among non-modifiable risk factors, one should mention gender and ethnicity as statistical data often shows correlation between those and cardiovascular diseases. In particular, prevalence of cardiovascular disease among white men by about 5.5% higher than that of women, which correlates well with relative prevalence of CVD normalized

by prevalence of physical inactivity for men and women [12] (See Table 2). However, the relative prevalence of CVD normalized by the prevalence of obesity is by 13% higher for white women than for men. Meanwhile the CVD prevalence among black men is about 8.1% lower than the CVD prevalence among black females. After normalization by either physical inactivity or obesity the relative prevalence for this family of diseases for black men is actually higher than for black women. This analysis shows that difference in gender doesn't probably play a significant role in development of cardiovascular disease after corrections for the primary risk factors like smoking, physical inactivity and obesity are made.

Table 2: Prevalence of Cardiovascular Disease among Different Ethnic Groups.

CVD Prevalence, Total and Relative to Risk Factor	Whites		Blacks		Mexican Americans	
	Males	Females	Males	Females	Males	Females
Total, %	34.3	32.4	41.1	44.7	29.2	29.3
Physical Inactivity	0.52	0.48	0.58	0.56	-	-
Obesity	0.49	0.57	0.65	0.58	0.39	0.41
Smoking	1.42	1.58	1.76	2.20	-	-

Similar to gender, ethnicity most probably has nor real correlation to cardiovascular disease. Indeed, although African Americans exhibit higher prevalence of CVD, their rural counterparts in Africa, which follow traditional low-fat and high-fiber diet, enjoy life without significant incidence of this disease. According to [45], "...coronary heart

disease (CHD) is near absent in rural areas” of Africa and is “... very uncommon in urban centers”. People in modern China, whose diet in many respects is similar to American traditional diet in terms of consumption of meat and animal fats has CVD mortality rate of about 40% which is close to 37% for Americans. Thus, we conclude that gender and ethnicity must be considered as derived or secondary risk factors, and, therefore, have to be omitted from the reverse risk assessment procedure once the primary risk factors are properly accounted for.

Summarizing the results, the relative risk for cardiovascular disease due to behavioral, non-diet related risk factors can be written as

$$(5) \chi = R_s R_p R_o$$

with the quantities R_s , R_p , and R_o given by formulae (1) – (3).

2.3.2 Prevalence of Smoking, Obesity, and Physical Inactivity

The previous section dealt with the relative risk of contracting CVD, which is associated with smoking, obesity and physical inactivity. To estimate impact of those risk factors on the probability of developing a chronic disease it is necessary to know the prevalence of those risk factors and their combinations among general population. Below we assume that the three aforesaid non-diet related risk factors are more or less independent from each other. In other words, a person which is physically inactive may not be necessarily a smoker or obese. By the same token, a smoker may not be obese and vice versa. On the other hand, those people, who have a combination of two risk factors or more, are obviously characterized by a higher probability to suffer from a chronic disease. Thus, it is essential to account for prevalence of all possible combinations of non-diet risk factors.

Table 3 summarizes statistical data on the prevalence of smoking, obesity and physical inactivity among Americans [12]. Each row in Table 3 corresponds to a

Table 3: Prevalence of Smoking, Obesity and Physical Inactivity.

No Risk Factors	Smoking	Overweight	Phys. Inactivity	Expression	Prevalence
+	-	-	-	Eq.(5)	0.194
-	+	±	±	π_s	0.209
-	±	+	±	π_o	0.651
-	±	±	+	π_p	0.301
-	+	+	+	$\pi_s \pi_o \pi_p$	0.041
-	+	-	-	$\pi_s (1-\pi_o)(1-\pi_p)$	0.051
-	-	+	-	$\pi_o (1-\pi_s)(1-\pi_p)$	0.360
-	-	-	+	$\pi_p (1-\pi_s)(1-\pi_o)$	0.083
-	+	+	-	$\pi_s \pi_o (1-\pi_p)$	0.095
-	+	-	+	$\pi_s \pi_p (1-\pi_o)$	0.021
-	-	+	+	$\pi_o \pi_p (1-\pi_s)$	0.155

combination of non-diet related risk factors. The plus sign indicates presence of a risk factor, while minus pertains to its absence. In addition to that, the plus/minus symbol corresponds to the situation where a risk factor may be either present or not. Quantities

π_s , π_o , π_p pertain to the prevalence of those people who smoke, are overweight and physically inactive, respectively. Their values have been obtained from statistical data [12]. All other values have been calculated by formulas shown in the next to the last column in Table 3.

The first row shows that the percentage of people, who don't smoke, are not obese and exercise regularly, is about 19.4%. This value was derived from expression (5), assuming that smoking, obesity and physical inactivity are independent risk factors.

$$(6) \pi_n = 1 - [\pi_o + \pi_s(1-\pi_o)(1-\pi_p) + \pi_p(1-\pi_s)(1-\pi_o) + \pi_s\pi_p(1-\pi_o)]$$

The meaning of equation (6) is quite clear: prevalence of people, who don't have non-diet risk factors, is equal to unity minus prevalence of four groups of people. The first of them comprises obese people. The second pertains to those who smoke, are not obese, and are physically active. The third group corresponds to people who physically inactive but are not obese and don't smoke. Finally, the fourth group includes those individuals who smoke and physically inactive, but are not obese.

The second, third and fourth rows in Table 3 corresponds to the prevalence of any single risk factor, irrespective to the presence of others. The fifth row contains plus sign for each of the three risk factors and pertains to the people who smoke, are obese and physically inactive simultaneously. The remaining rows describe all other possible combinations. From Table 3, it is seen that the percentage of those people who are overweight or obese is pretty high and significantly affects all other groups.

Assuming that the probability of developing a cardiovascular disease is proportional to the relative risk times prevalence, we find the following expression for the total relative risk associated with prevalence of the non-diet related risk factors among American population

$$(7) R_{\text{non-diet}} = \pi_n + \pi_s (1-\pi_o)(1-\pi_p) \langle R_s \rangle + \pi_o (1-\pi_s)(1-\pi_p) \langle R_o \rangle + \\ \pi_p (1-\pi_s)(1-\pi_o) \langle R_p \rangle + \pi_s \pi_p (1-\pi_o) \langle R_s \rangle \langle R_p \rangle + \pi_s \pi_o (1-\pi_p) \langle R_s \rangle \langle R_o \rangle + \\ \pi_o \pi_p (1-\pi_s) \langle R_o \rangle \langle R_p \rangle + \pi_s \pi_o \pi_p \langle R_s \rangle \langle R_o \rangle \langle R_p \rangle,$$

where $\langle R_s \rangle$, $\langle R_o \rangle$, and $\langle R_p \rangle$ denote the average relative risk associated with smoking, being overweight and physically inactive respectively.

The actual value of the quantity $R_{\text{non-diet}}$ can be calculated from formula (6) by substituting the average relative risk values for smoking, being overweight and physically inactive. Those values are derived from Eqs. (1) – (3), based on the statistical data from literature [46-48]. Thus, the value of $\langle R_s \rangle = 2.125$ is obtained assuming the average number of 15 cigarettes per day. The average relative risk due to being overweight is estimated as $\langle R_o \rangle = 1.492$, since the average BMI for overweight people in the U.S.A. is about 28.2 [49, 50]. When calculating the average relative risk of physical inactivity, we assumed that physically inactive person spends about 1.25 hours per week on physical activity during the leisure time. According to formula (2), this results in the value $\langle R_p \rangle = 1.5$. Substitution of the average relative risks into formula (6) yields the following value

for the total relative risk of non-diet related factors, associated with the prevalence of smoking, obesity and physical inactivity in the United States:

$$(7) R_{\text{non-diet}} = 1.87$$

The latter result shows that prevalence of smoking, being overweight and physically inactive significantly influences the probability of developing cardiovascular disease by almost doubling the value of the relative risk, which would be equal to unity otherwise.

2.3.3 Cumulative Effect of Exposure to Diet-Related Risk Factors

Recent scientific studies [12, 51-55] indicate that diet plays an important role in an overall healthy lifestyle. In particular, exposure to hazardous substances (risk factors), which are brought into a human body with food, has been recognized as a major contributor to the risk of contracting a chronic disease such as cancer, cardiovascular or liver disease. The risk factors associated with contracting a certain condition and coming with daily food intakes are called dietary risk factors and normally include (a) saturated fat, (b) trans fat, (c) cholesterol, (d) table salt, (e) alcohol, etc. The risk is generally characterized by the probability for an event (contracting a disease) to occur and the impact of this event on the human being. The impact of contracting a chronic disease on the human body and health may be quite severe and is no longer being questioned. Indeed, in spite of major advances in modern medicine, maintaining a healthy diet offers the greatest potential of all known approaches for reducing the risk of cardiovascular and

other chronic diseases [1]. Therefore, one the major tasks in a reverse risk assessment problem is to estimate the probability of contracting the condition as a function of time and exposure to the risk factors.

Let us assume that there are n independent diet risk factors r_i ($i=1,2,\dots,n$), which contribute to development of a chronic disease. We denote $p(t)$ the probability of contracting the disease at a certain age t by an average person in a large population group. The probability $p(t)$ is assumed to be proportional to the total exposure to the risk factors during the period of time t . Since the exposure to the risk factors has a cumulative effect, the change in the quantity p within a small time dt can be written, without losing generality, as a sum over partial probabilities dp_i

$$(1) \quad dp = [1-p(t, \mathbf{E})] \sum_{i=1}^n dp_i = [1-p(t, \mathbf{E})] \sum_{i=1}^n k_i(t) dE_i ,$$

where dE_i is the differential risk exposure associated with the risk factor r_i , $k_i(t)$ is a proportionality coefficient, which accounts for the influence of the environmental changes on the partial probability dp_i to contract the disease. In formula (1), summation is performed over all diet-related risk factors n . The multiplier $[1 - p]$ in the right-hand side of Eq.(1) describes the probability to find a healthy person and accounts for the trivial normalization condition. The infinitesimally small exposure dE_i is determined by the hazardous substance intake I_i times the exposure interval dt . Thus, we can present the exposure as

$$(2) \quad dE_i = I_i(t) dt ,$$

After substitution of Eq. (2) into expression 1 we arrive the formula, which relates the differential probability dP and the intakes $I_i(t)$ ($i=1,2,\dots,n$)

$$(3) \quad dp/(1-p) = \sum_{i=1}^n k_i(t) I_i(t) dt,$$

Assuming that without exposure to risk factors the probability of contracting the disease is zero and performing integration of the both sides of Eq. (3), we arrive at the expression for the probability p in the form of a sum over all intakes integrated over time from 0 to t :

$$(4) \quad p(t) = 1 - \exp\left(- \sum_{i=1}^n \int_0^t k_i(t') I_i(t') dt'\right), \quad (0 < p(t) < 1).$$

By its physical meaning, expression (4) describes the cumulative effect of each of the risk factors involved. As it follows from Eq.(4) in the limiting case of small exposures, the contribution of the risk factor r_i is presented by the partial probability

$$(5) \quad p_i(t) = \int_0^t k_i(t') I_i(t') dt', \quad (0 < p_i(t) \ll 1),$$

Thus, the quantity $p_i(t)$ is associated with the intake I_i integrated over the period of time from 0 to t and describes the evolution of probability to contract the disease due to the risk factor r_i in the limiting case of small times $t \ll T$, where T is the lifespan. It should be noted that expression (4) is pretty general and can be also used to account for influence of such non-dietary risk factors like smoking and physical inactivity.

Eqs. (4) and (5) contain environmental factors $k_i(t)$ as functions of time, which accounts for individual total calorie intake needs and are not known in advance. Those functions should be determined from either experiment or statistical data. For simplicity, we will confine our further analysis to the limiting case of large times t , which obey the condition

$$(6) \quad t \gg \tau,$$

where τ is a typical intake time. Approximation (6) means essentially that many intakes of hazardous substances take place before a noticeable change in the health condition of an individual becomes apparent. Secondly, we will focus on determining the risk of contracting a disease at the time t comparable with the average lifespan T , $t \sim T$.

Since the condition $T \gg \tau$ usually holds true, we replace the upper limit of integration in Eq.(4) by infinity ∞ and obtain the formula for the probability P of contracting a chronic disease

$$(7) \quad p(T) = 1 - \exp\left(- \sum_{i=1}^n \int_0^{\infty} k_i(t') I_i(t') dt'\right),$$

To further simplify Eq.(7) we apply the steady-state approximation by replacing the coefficients k_i and intakes I_i by their time-averages. As a result, formula (7) can be rewritten as a sum of products of the average intakes I_i and coefficients k_i

$$(8) \quad p = 1 - \exp\left(- T \sum_{i=1}^n k_i I_i\right),$$

where the quantities k_i and I_i are given by the expressions

$$(9) \quad k_i = T^{-1} \int_0^{\infty} k_i(t') dt' ,$$

and

$$(10) \quad I_i = T^{-1} \int_0^{\infty} I_i(t') dt' ,$$

Approximation (8) is substantiated if the functions $k_i(t')$ or $I_i(t')$ are slowly varying in time. In particular, it holds true if the intake pattern doesn't change significantly over the individual lifetime. If the lifespan T is measured in days, the quantity I_i represents an average daily intake of the hazardous substance i , whereas the coefficient k_i is proportional to the average value of the function $k_i(t)$. In the limiting case of small influence of dietary factors on developing the disease, expression (8) reduces to

$$(11) \quad p = T \sum_{i=1}^n k_i I_i ,$$

Hence, in this limiting case the probability p becomes a linear function of the average daily intakes. Expression (11) also shows the cumulative effect of diet-related risk factors: the larger number of risk factors is involved the higher is the probability to develop a chronic disease.

2.3.4 Simple Formula for the Probability of Developing a Chronic Disease

The linear dependence of the probability to develop a chronic disease on the daily intakes and the obvious observation that in absence of hazardous-substance intakes the risk is supposed to be zero, prompt an alternative approach of deriving an expression for

probability p . This approach is based on expanding the probability to contract a chronic disease by the end of lifespan (on average, after the age of 70 – 75 years) into a Taylor series near the zero point. Introducing the intake vector

$$(12) \quad \mathbf{I} = \{ I_1, I_2, \dots, I_n \},$$

we can present the probability $p = p(\mathbf{I})$ is a function of the vector \mathbf{I} in the multidimensional space of the hazardous-substance intakes. Expanding the quantity $p(\mathbf{I})$ in the vicinity of $\mathbf{I} = 0$, we find

$$(13) \quad p(\mathbf{I}) = p(0) + \text{grad } P(0) \cdot \mathbf{I} + (1/2) \sum [\partial^2 p(0) / \partial I_i \partial I_j] I_i I_j + \dots$$

We note that that $p(0) = 0$ and, therefore, the first term in the right-hand side of Eq. (12) can be omitted. Assuming further that quadratic and other higher order terms are small compared to the linear ones, we can retain only the linear terms in the right-hand side of expression (13). This significantly simplifies the expression for the probability P_d , which now can be written as

$$(14) \quad p(\mathbf{I}) = \text{grad } p(0) \cdot \mathbf{I}.$$

Formula (14) relates the probability p with its gradient over \mathbf{I} in the vicinity of the zero point $\mathbf{I} = 0$ with daily intakes as components of the intake vector \mathbf{I} . Similar to Eq.(13), formula (14) establishes a linear relationship between the probability of contracting a

disease and daily intakes. However, the advantage of Eq.(14) is that the right hand-side explicitly depends only on the daily intakes and first derivatives of the probability p . The latter can be obtained from statistical data available in literature [12, 34, 53].

Let us denote q the probability to develop a chronic condition due to smoking, being overweight or physically inactive. Applying the same considerations and assuming that in the first approximation the relative risk associated with smoking, obesity and physical inactivity is a linear function of daily exposure to those hazards, we find the following expression for the quantity q :

$$(15) \quad q(\mathbf{K}) = \text{grad } q(0) \cdot \mathbf{K} ,$$

where \mathbf{K} is the daily hazard exposure vector.

The total probability P to develop a chronic disease is a sum of quantities p and q :

$$(16) \quad P = p + q$$

Substitution of formula (14) and (15) into (16) yields:

$$(17) \quad P(\mathbf{I}, \mathbf{K}) = \text{grad } p(0) \cdot \mathbf{I} + \text{grad } q(0) \cdot \mathbf{K}$$

In the sections below, formula (17) is applied to estimation of the probability to develop a cardiovascular disease due to both diet-related and behavioral risk factors.

2.3.5 Diet-Related Probability of Developing Cardiovascular Disease

As mentioned in Chapter 1, cardiovascular disease (CVD) represents a family of diseases associated with deterioration of the cardiovascular system. The importance of CVD prevention can be hardly overestimated as since 1900 this disease has been the No. 1 killer in the United States. The huge amount of statistical data accumulated for CVD [12] allows

one to estimate contribution of different risk factors to the probability of developing the disease.

Table 4 shows differential probabilities for developing CVD due to intake of hazardous substances like saturated fat, cholesterol, trans fat, salt and alcohol at a later age in the United States. The respective values have been derived by formula (14) for 2000 kCal diet and the LDL blood concentration of 100 mg/dl. In particular, the differential probability for saturated fat was obtained assuming that reduction in 1% of saturated fat intake results in reducing the LDL level by 1.5 mg/dl or 1.5%. This reduction, in turn, is equivalent to reducing the risk of CVD events by 1.5% [15]. Similarly, cholesterol intake of 0.1 g augments the LDL level by 2 mg/dl, which corresponds to CVD risk increase of 2% [15]. The differential probability for trans fat is estimated based on experimental results which show that increase in trans fat intake from 1.3 g to 6.2 g among women increase, on average, the risk of CVD by 27% [15].

Table 4: Average Daily Consumption of Hazardous Substances.

Risk Factors	Saturated Fat	Cholesterol	Trans Fat	Sodium
Probability Gradient, g⁻¹	0.0068	0.20	0.055	0.026
Daily Consumption, g	25	0.290	2.8	3.6
Partial Probability	0.17	0.058	0.154	0.094
Total Probability p	0.476			

2.3.6 Effect of Smoking, Obesity and Physical Inactivity

Expression (17) describes the probability P as a function of daily exposure to hazards in the linear approximation over two groups of risk factors and is valid under the assumption that the risk factors involved are not correlated.

To determine contribution of non-diet risk factors to the probability P we note that the partial non-diet risk probability q_i is equal to the first derivative of the probability q on the daily hazard exposure K_i times the exposure K_i , $q = (\partial q / \partial K_i) K_i$. On the other hand, the quantity q_i can be calculated from the relative risk associated with the i -th non-diet risk factor by using Eq(1). Knowledge of the partial probability q_i and the average daily exposures to no-diet related hazards, which are available in literature, allows one to calculate the values for the gradient $\text{grad } q(0)$. Substitution of those values and the probability p from Table 3 into Eq(1) yields non-behavioral risk gradient components. The respective results obtained by formula (15) are shown in Table 5.

The smoking-related parameters have been calculated based on the assumption that the relative risk to develop a CVD due to smoking is on average twice as high as compared to those people who are not exposed to smoking, are not overweight and are physically active. This assumption is reasonable and corresponds to the lowest value of smoking-related risk which ranges from 2 to 4 [15]. It has been also assumed that the average number of cigarettes smoked per day is 16.8 [55]. The probability gradients for being overweight and physically inactive have been calculated assuming that relative risks associated with both of those risk factors is on average 1.5 for general population. The average BMI value was set to 29.2 since two thirds of overweight people have BMI greater

than 25 but less than 30, while one third has BMI greater than 30. Finally, the average number of hours spent on physical activity during leisure time by those people who are considered to be inactive was set to 1.25 per week, which half the threshold for physical inactivity.

As it follows from Table 5, the average exposure to smoking, physical inactivity and being overweight combined together make the probability for contracting a cardiovascular disease close to unity $q \sim 1$. Thus, an average person which is exposed to all three non-diet risk factors and follows traditional American diet will almost unavoidably develop CVD by the end of his/her life.

Table 5: Average Daily Exposure to Non-Diet Hazards and CVD Probability.

Risk Factors	Smoking	Overweight	Physical Inactivity
Probability Gradient	0.028, cig ⁻¹	0.0082 m ² /kg	0.00317, week/min
Daily Exposure	16.8 cigarettes	29.2 kg/m ⁻²	75, min/week
Partial Probability	0.476	0.239	0.239
Total Probability q	0.952		

2.3.7 Differential Probabilities for Developing Cardiovascular Diseases

In practice, it is often necessary to determine the probability for developing a particular condition which belongs to the family of cardiovascular diseases. All CVD can be conveniently divided into three large groups: (1) Heart Disease, (2) Stroke and (3) Hypertension. Although, there are no statistical data on a differential probability to develop a condition due to a certain risk factor, it is reasonable to assume that the relative prevalence of heart disease, stroke and hypertension is a good indicator of this kind of risk and respective probability. In Table 6, the absolute statistical data on the prevalence of cardiovascular disease in 2003 [12] are used to estimate the relative probability for an average American to develop heart disease, stroke or hypertension. Note that heart disease includes both coronary heart disease (13,200,000) and heart failure (5,000,000). As it

Table 6: Prevalence of Cardiovascular Diseases (2003).

Disease	CVD	Heart Disease	Stroke	Hypertension
Prevalence	71,300,000	18,200,000	5,500,000	65,000,000
Relative Probability	1	0.255	0.077	0.912

follows from Table 6, a person with CVD will almost certainly have hypertension. On the other hand, the risk of developing stroke remains relatively small as compared to hypertension and heart disease.

2.4 Analysis Class Diagrams

The mathematical model provides an important analytical foundation for the risk assessment tool. It is still necessary to perform analysis of the user interaction with the software to be designed. In this section, we consider the analysis class diagrams pertaining to the basic use cases outlined earlier. The analysis model is supposed to describe what the risk assessment tool does in terms of traditional boundary, control and entity objects. Figs. 12 – 21 illustrate collaboration diagrams for use case realization in the analysis model.

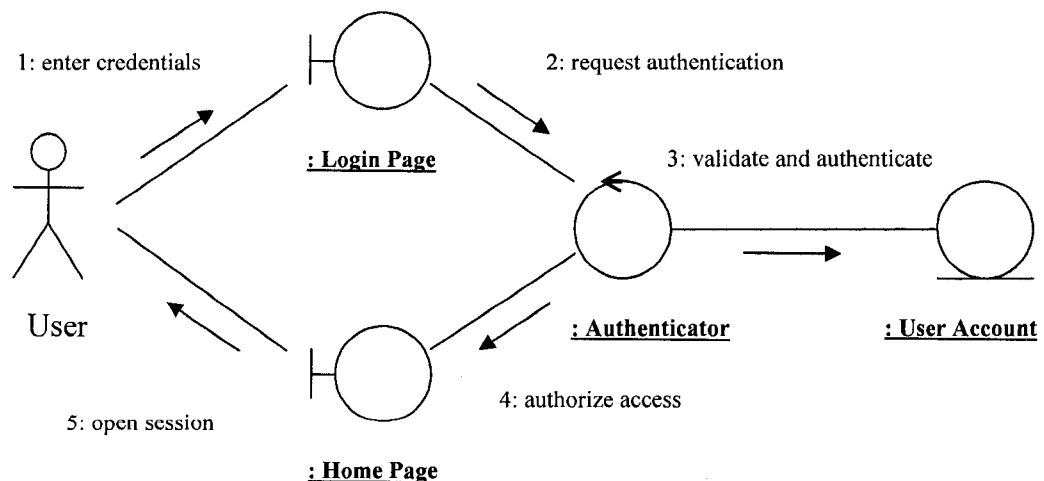


Figure 12: Collaboration Diagram for “Login” Use Case.

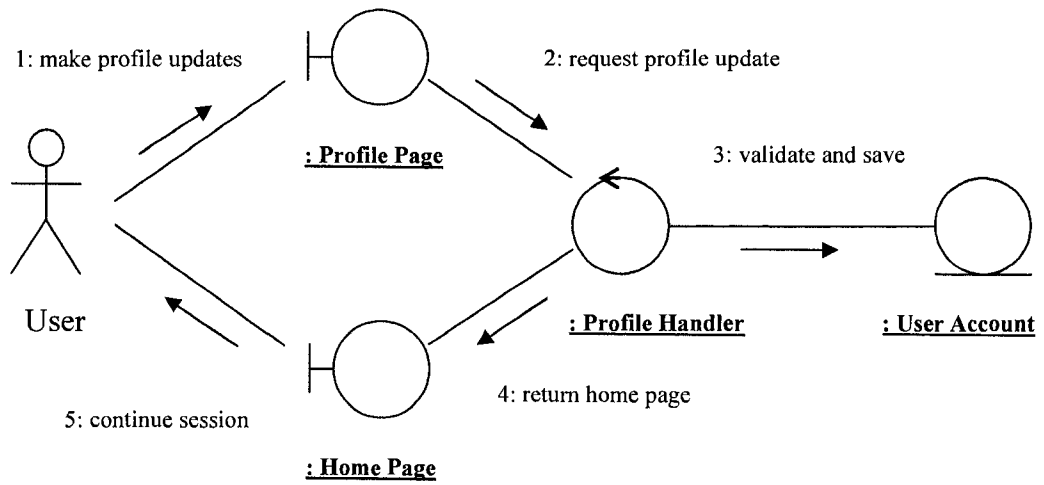


Figure 13: Collaboration Diagram for “Update Profile” Use Case.

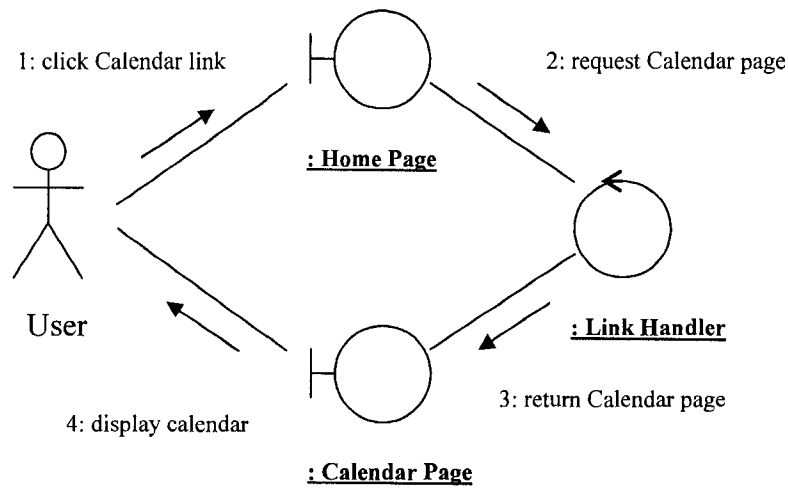


Figure 14: Collaboration Diagram for “Look up Calendar” Use Case.

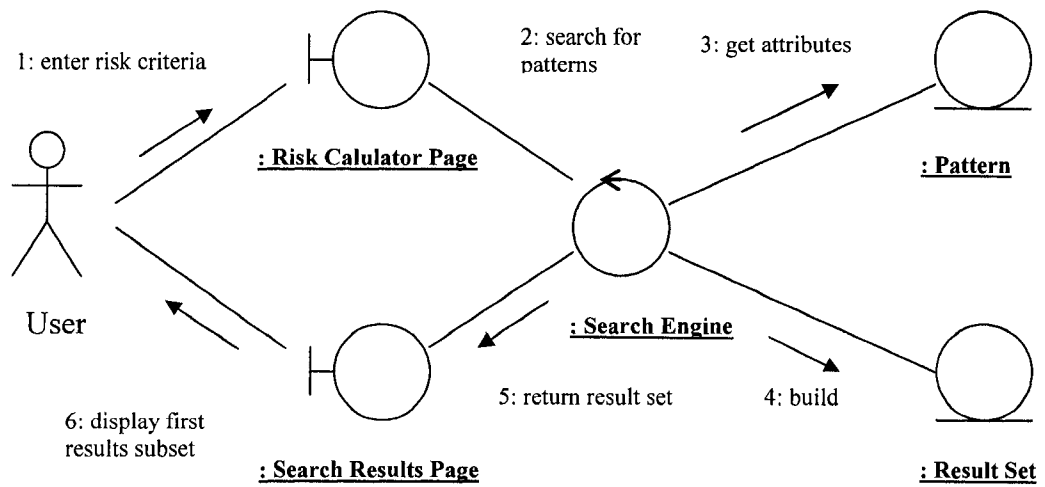


Figure 15: Collaboration Diagram for “Calculate Risk” Use Case.

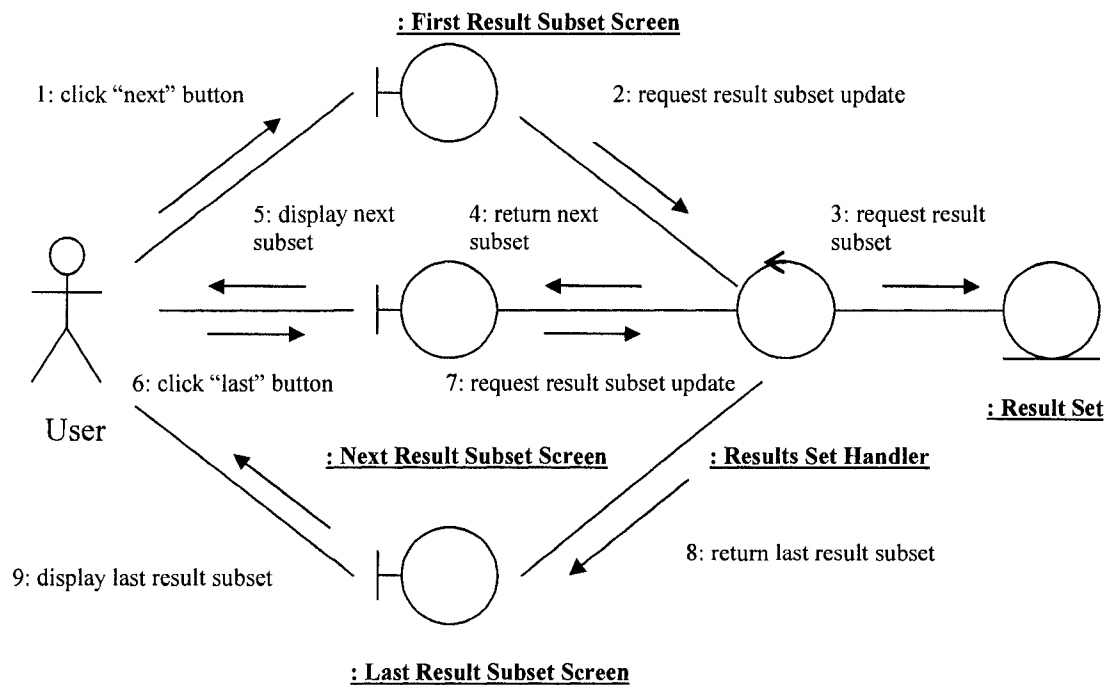


Figure 16: Collaboration Diagram for “Browse Results” Use Case.

It should be noted that for the sake of clarity, the user action to retrieve the previous result subset is omitted in Fig. 16. For the same considerations, the user action to retrieve the “Minerals” data in the food intake pattern is omitted in Fig. 17. Those simplifications, however, are trivial and don’t diminish the meaning of the respective collaboration diagrams. Also, the “K” letter link is used in Fig. 18 as an example for searching an entry in the glossary, which starts with a letter “K”.

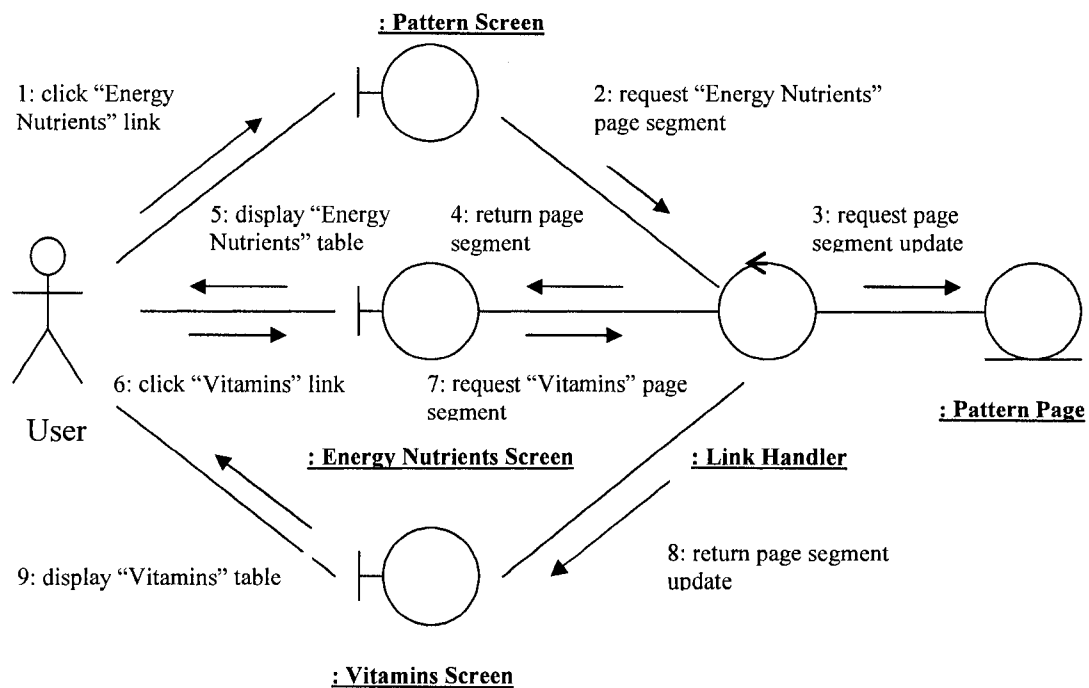


Figure 17: Collaboration Diagram for “Explore Pattern” Use Case.

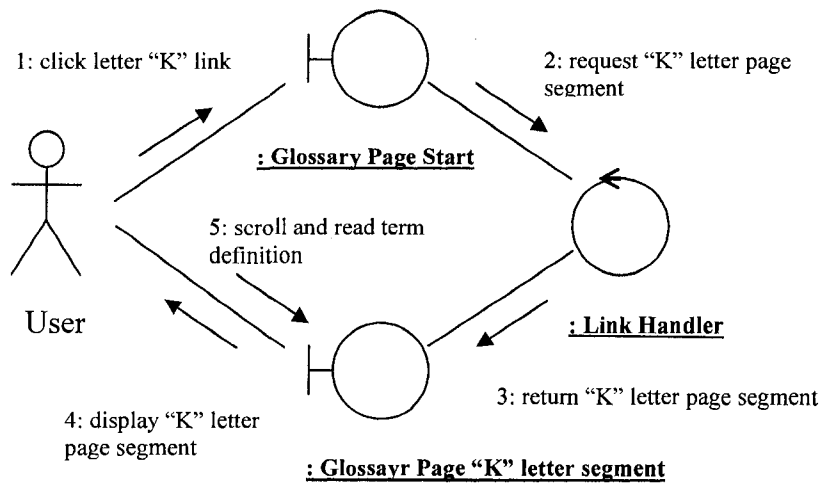


Figure 18: Collaboration Diagram for “Look up Glossary” Use Case.

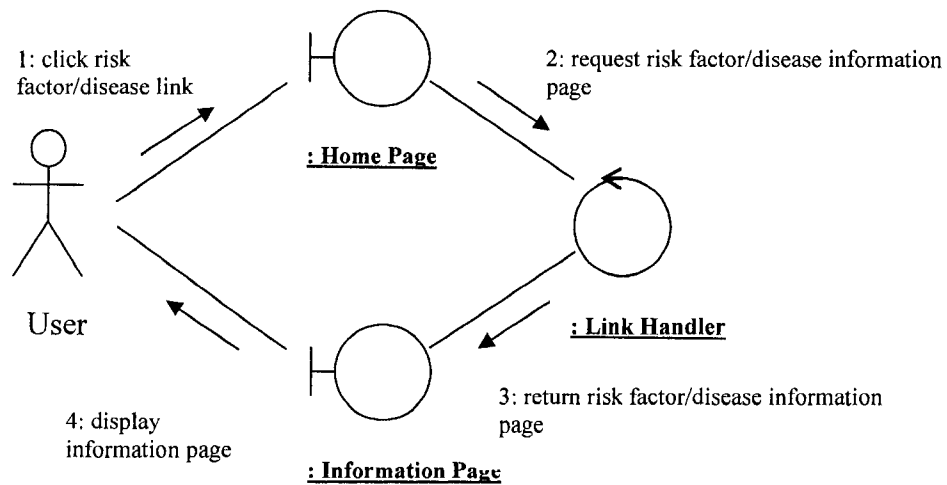


Figure 19: Collaboration Diagram for “Look up Calendar” Use Case.

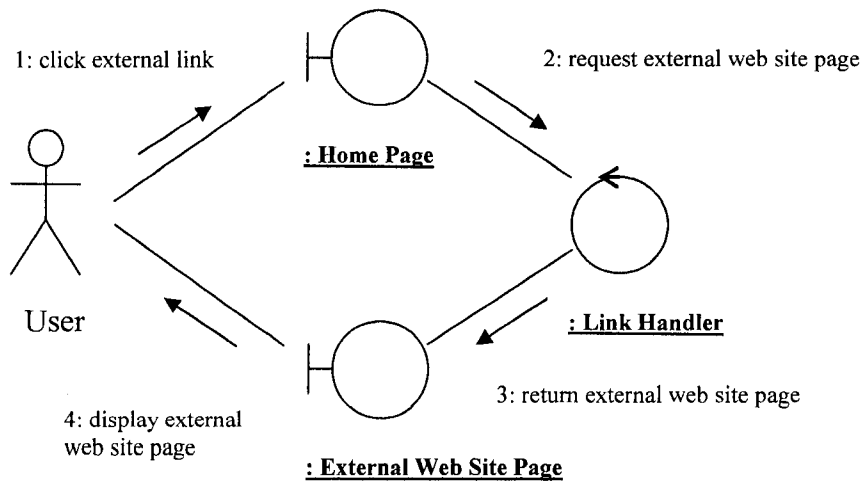


Figure 20: Collaboration Diagram for “Use External Link” Use Case.

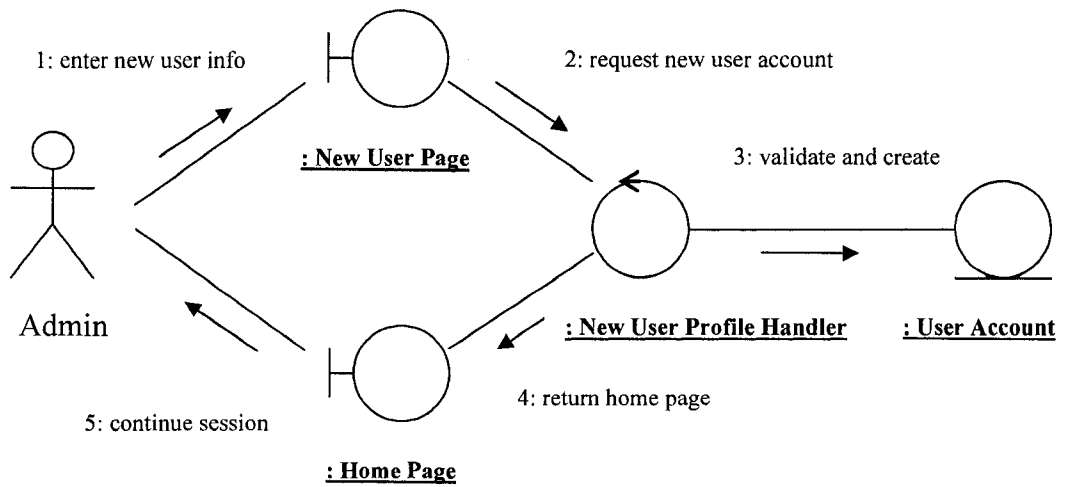


Figure 21: Collaboration Diagram for “Create New Profile” Use Case.

2.5 Web Page Architecture and Navigation

As it follows from the previous section, the user interaction with the application is focused on risk calculations and access to related resources. This prompts the web architecture which allows the user to freely navigate from one informational page to another, being able at the same time to use the risk calculator utility from almost any application page. According to the system requirements, the informational pages are supposed to include information about diet-related risk factors (cholesterol, fat, sugar, alcohol, salt, fiber, vitamins, and mineral) and diet-related chronic diseases (heart disease, stroke, hypertension, diabetes, obesity, osteoporosis, diverticulosis, and dental caries). The web page architecture and navigation are schematically shown on Fig. 22.

For the sake of brevity, Fig. 22 doesn't show all connections among different pages. There are 39 pages and an attempt to display all links among those would require more than one thousand connections. Instead, all the pages are conveniently split into four groups, based on their link status. Thus, all the pages which are directly connected between each other are shown as solid rectangles. Those include "Home", "Risk Calculator", "Glossary", "Profile", and informational pages of the first level, containing information about risk factors and diseases. The second group of pages shown as dashed rectangles includes pages which are directly accessible only from their parent. This group comprises all FAQ (Frequently Asked Questions) pages related to risk factors and diseases and also "Search Results" page. The latter is invoked by a link from "Risk Calculator" page. The third and the fourth page groups consist just of one page "Login"

and “New User”, respectively. The “Login” page shown as dashed, double-dotted rectangle has a special status. It is a starting page providing entry into the application.

Upon login the user is presented with the “Home” page from which he/she can navigate to a page of choice. Finally, the “New User” page shown by dotted rectangle is available only from the “Profile” page and only to the user with the administrative privileges. In other words, only a system administrator is supposed to create new user profiles.

2.6 Data Model

From the discussion in Chapter 1, it follows that a reverse risk assessment system, in general, should include at least two components. One of them is supposed to provide the user with the ability to make risk assessments while another should be able to generate, or otherwise simulate, a large number of possible scenarios associated with a particular outcome. In our case of diet-related chronic diseases, the first component would manage calculation logic for contracting a disease whereas the second component must generate various food intake patterns. Correspondingly, the data model of the system can be conveniently split into two pieces: (a) risk factors and chronic diseases, and (b) food intake patterns and their nutrient composition. In the next section the emphasis is put on analyzing major entities and their relationships in the considered risk assessment problem.

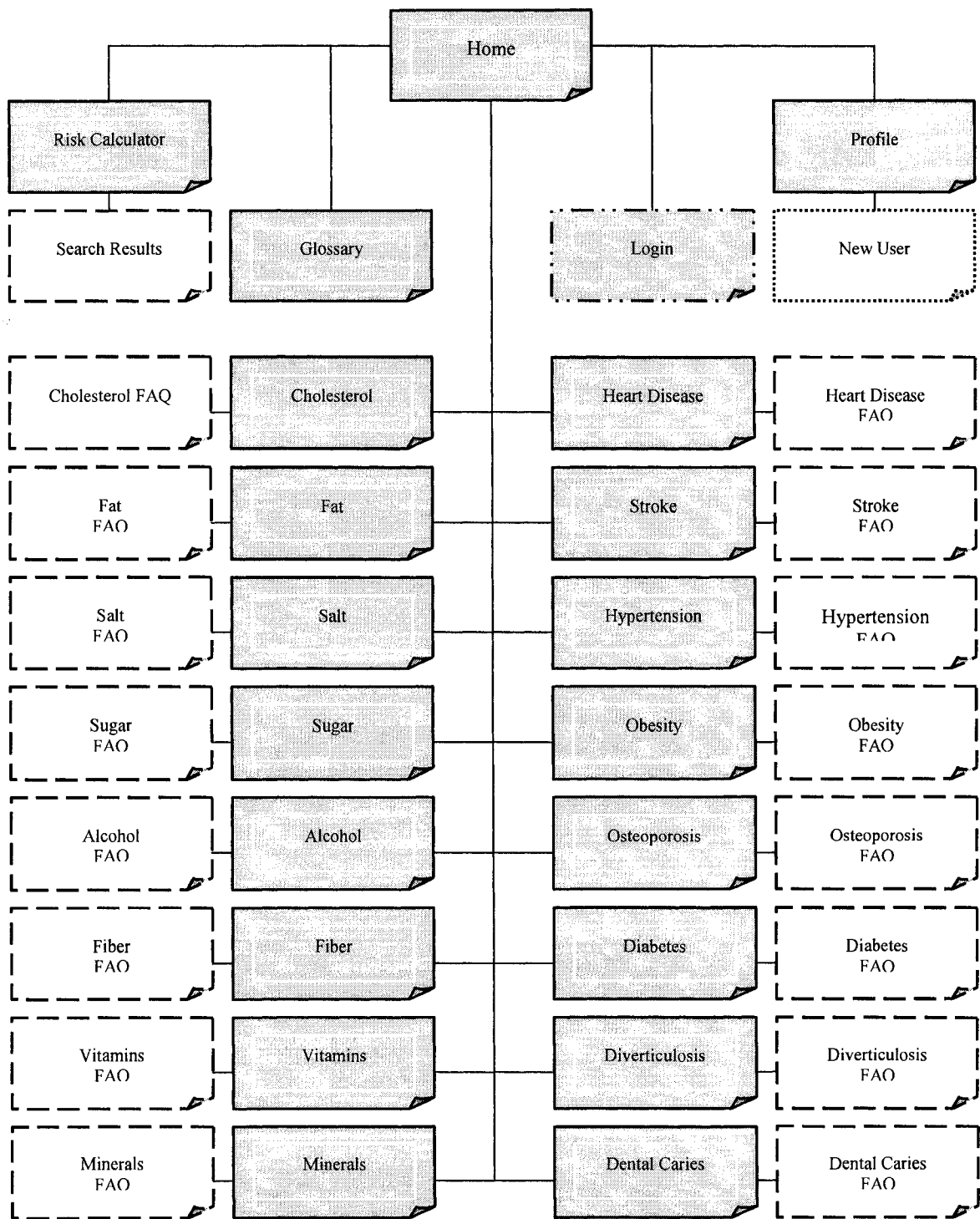


Figure 22: Web Page Architecture and Navigation.

2.6.1 Entity Relationships

The entity relationship analysis is an important step in designing and implementing the system database. Fig. 23 shows the entity-relationship diagram (ERD) for diet-related risk factors and chronic diseases. Obviously, the diagram can be easily generalized to include behavioral risk factors like smoking, physical inactivity, etc. However, these risk factors are deliberately omitted from the diagram since computationally they are accounted for as correction coefficients rather than properties of persistent entity objects. It should be noted that some of the chronic disease can exacerbate other ones and vice versa and, therefore, may themselves become risk factors. This, essentially non-linear, effect is not accounted for in the model presented here.

Similarly, Fig. 24 depicts the entity-relationship diagram for the second group of the most important entities which include food intake pattern, food intake, and nutrients. For the sake of clarity various kinds of minerals, vitamins and carbohydrates are not shown there. The bridge between diet-related risk entities and food intake pattern entities is provided by the relationship between the food intake and the risk factor (Fig. 25). To avoid many-to-many relationship a bridge entity “RiskFactorOccurrence” is introduced. The entity “RiskFactorOccurrence” describes presence of various diet-related risk factors in different food intakes.

As mentioned earlier, the nutrient structure in Fig. 25 is not shown for the sake of clarity. The full outline of the nutrient class, which is used for nutritional characterization of food intake patterns, is presented in Fig. 26. Entity-relationship diagrams in Figs. 25

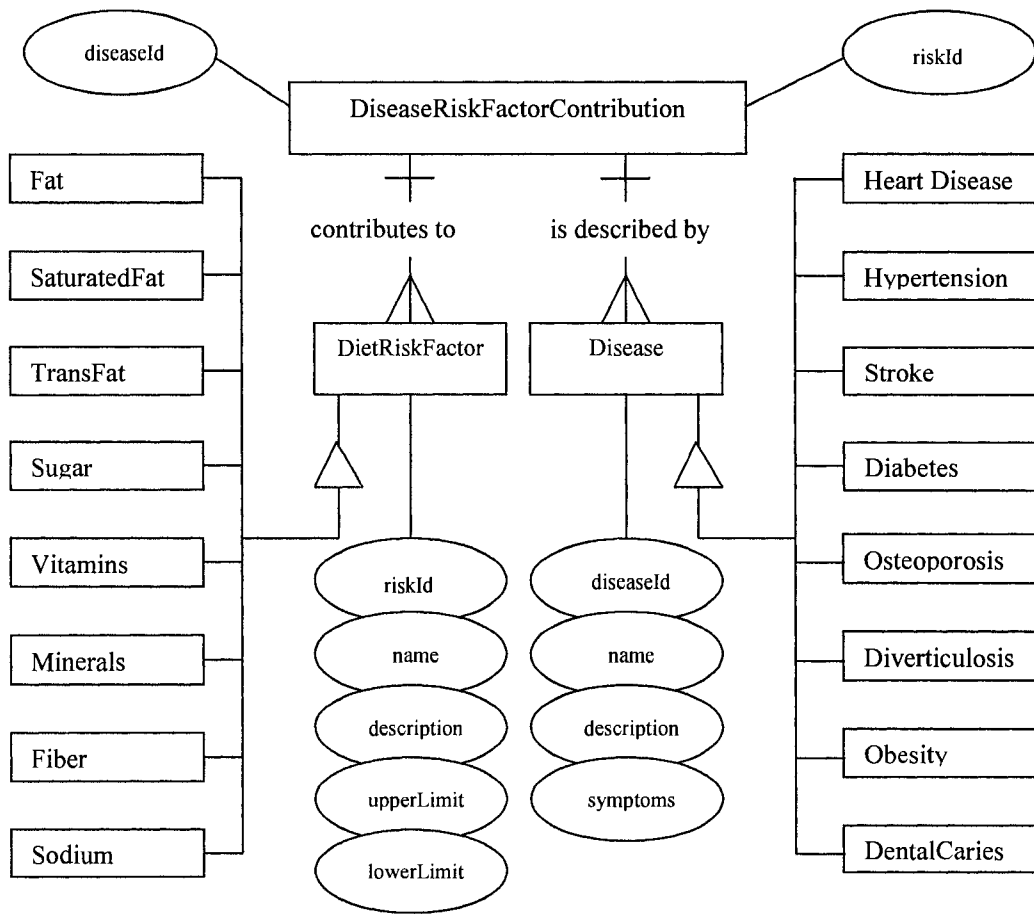


Figure 23: Diet-Related Risk Factor–Chronic Disease ERD.

and 26 are used to build database schema which is discussed in the next section.

2.6.2 Database Design

In view of relatively large numbers of essential nutrients (~ 40) in any food item and food items (20 – 30) in any statistically significant food intake pattern, characterization of a food intake involves knowledge of thousands of parameters. Under these circumstances

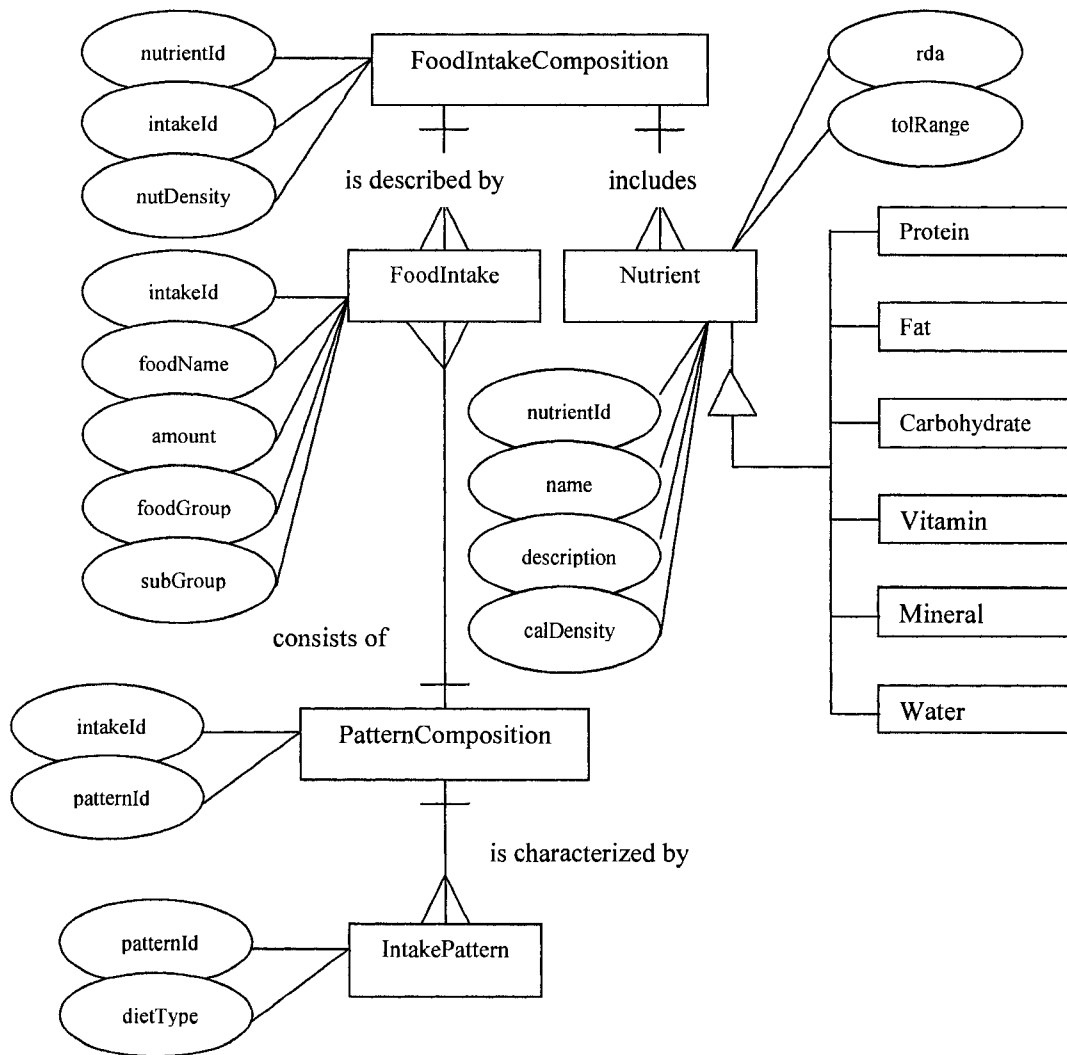


Figure 24: Food Intake Pattern ERD.

building a relational database schema from diagrams shown in Figs. 23 – 26 doesn't seem practical. Instead, it is suggested to create a simple two-table relational database and a database of XML files. One of the tables in the relational database is supposed to contain user profiles, while another table would be populated with pattern metadata which include 6–7 major pattern properties.

The detailed nutritional and disease risk related information is stored in the XML database of patterns. Each pattern file encapsulates all pattern characteristics in terms of its disease risk and nutritional value. Thus, a pattern file contains information about total amounts of nutrients in each food item of the pattern. In addition to that, each pattern file provides information about recommended daily intakes of the vitamins, minerals and major nutrients, so that the user could quickly evaluate the nutritional value of a particular food intake.

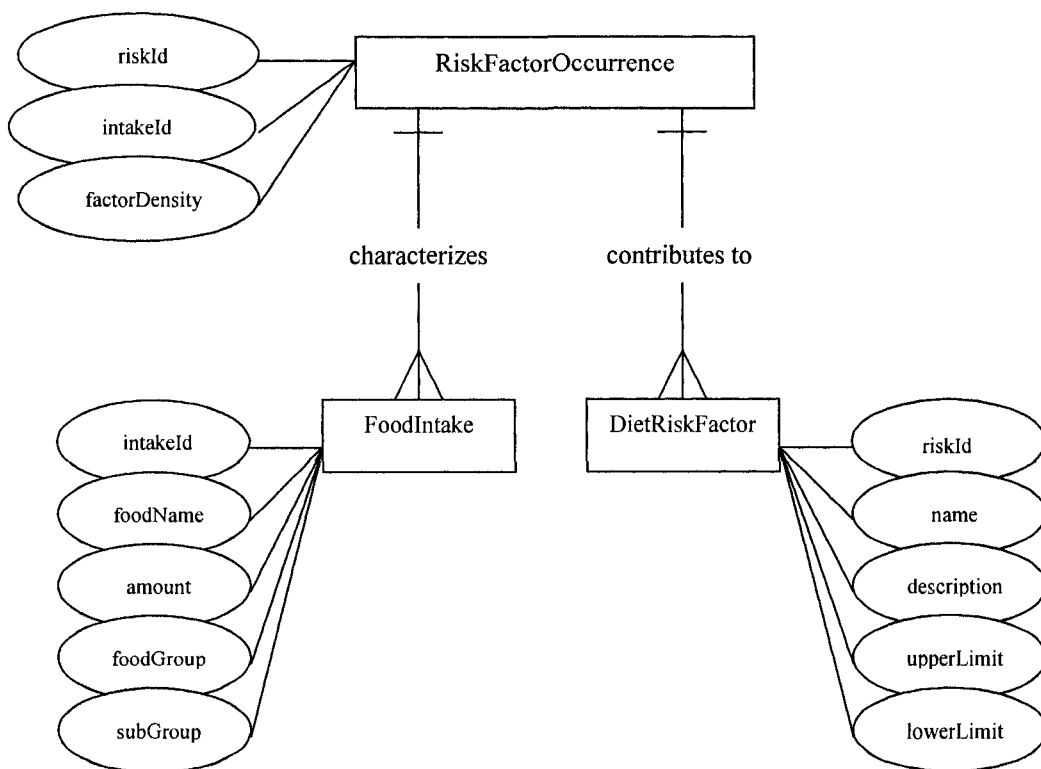


Figure 25: Bridge ERD (Food Intake Pattern and Risk Factor Entities).

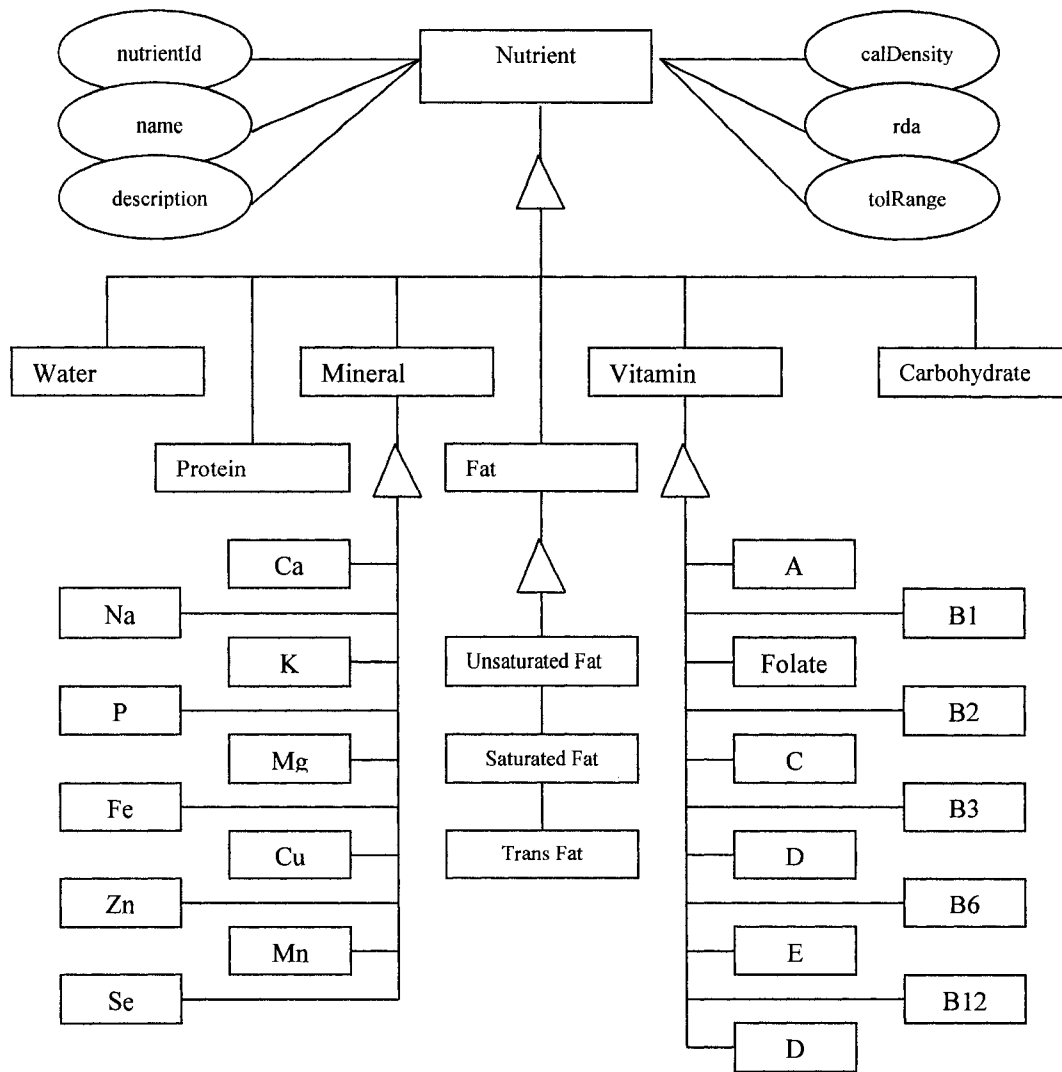


Figure 26: Nutrient Class Structure of Food Intake Patterns.

patternID	dietType	cvdRisk	energy	foodAmount
VARCHAR(11)	VARCHAR(30)	NUMBER(3)	NUMBER(5)	NUMBER(5)
)				

fat	sodium	satFat	filePath	summary
NUMBER(5)	NUMBER(6)	NUMBER(5)	VARCHAR(50)	VARCHAR(400)

Figure 27: The “Pattern” Table Schema.

Fig. 27 shows the structure of the main table “Pattern”. The table contains eleven columns with the patternID as a primary key. Columns “fat”, “satFat” and “sodium” contain data about content of the most significant risk factors in a particular food intake pattern. The columns “dietType”, “energy”, and “foodAmount” provide general characteristics of the pattern in terms of diet preference, calorie intake and total food amount, respectively. The overall cardiovascular disease probability associated with the respective food intake is provided in the column “cvdRisk”. The “summary” column is intended to display brief description of the food items which constitute the intake. This column is especially useful for searching the database when specifying food preferences. Finally, the column “filePath” contains information about the pattern file path in the XML database.

The second table in the relational database is dedicated for storing user profiles. The table structure is shown in Fig 28.

userID	firstName	lastName	email
VARCHAR(11)	VARCHAR(30)	VARCHAR(30)	VARCHAR(30)

role	password
VARCHAR(10)	VARCHAR(10)

Figure 28: The “User” Table Schema.

The table “User” contains seven columns with the userID as a primary key. Columns “firstName”, “lastName”, “email” and “password” are pretty much self-explanatory. The column role serves to keep information about the user role as related to the risk assessment application. For example, the user with the “admin” role may have access to features which are not available to the ordinary users.

The organization of the XML database of food intake patterns is shown schematically in Fig. 29. The root folder “Patterns” contains four subfolders corresponding to 4 different diet preferences: (a) Traditional American, (b) Mediterranean, (c) Vegetarian, and (d) Vegan. Each of those subfolders has the identical subfolder hierarchy with the first layer of subfolders pertaining to different degrees of cardiovascular disease risk and the second layer of folders subdividing food intake pattern files according to their calorie (energy) intake range. The “energy range” subfolders store the actual pattern files. A food intake pattern file has a strictly defined structure following the layout presented in Fig. 30. Each pattern file contains the main table named “Food Intake Description” with the information about different food items,

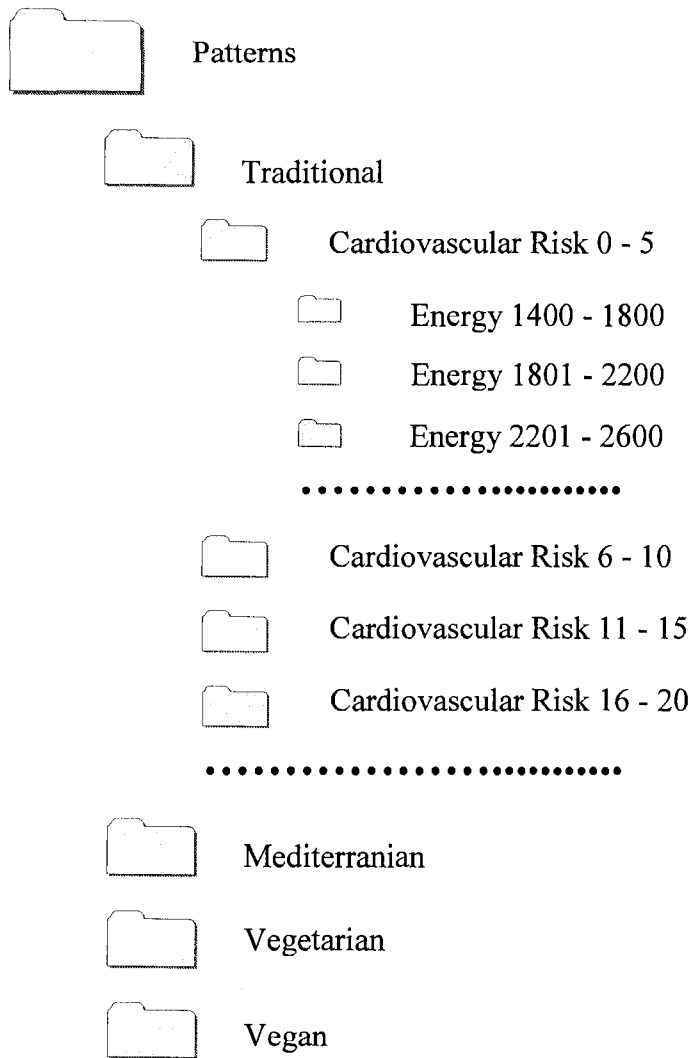


Figure 29: The Folder Structure of the XML Database.

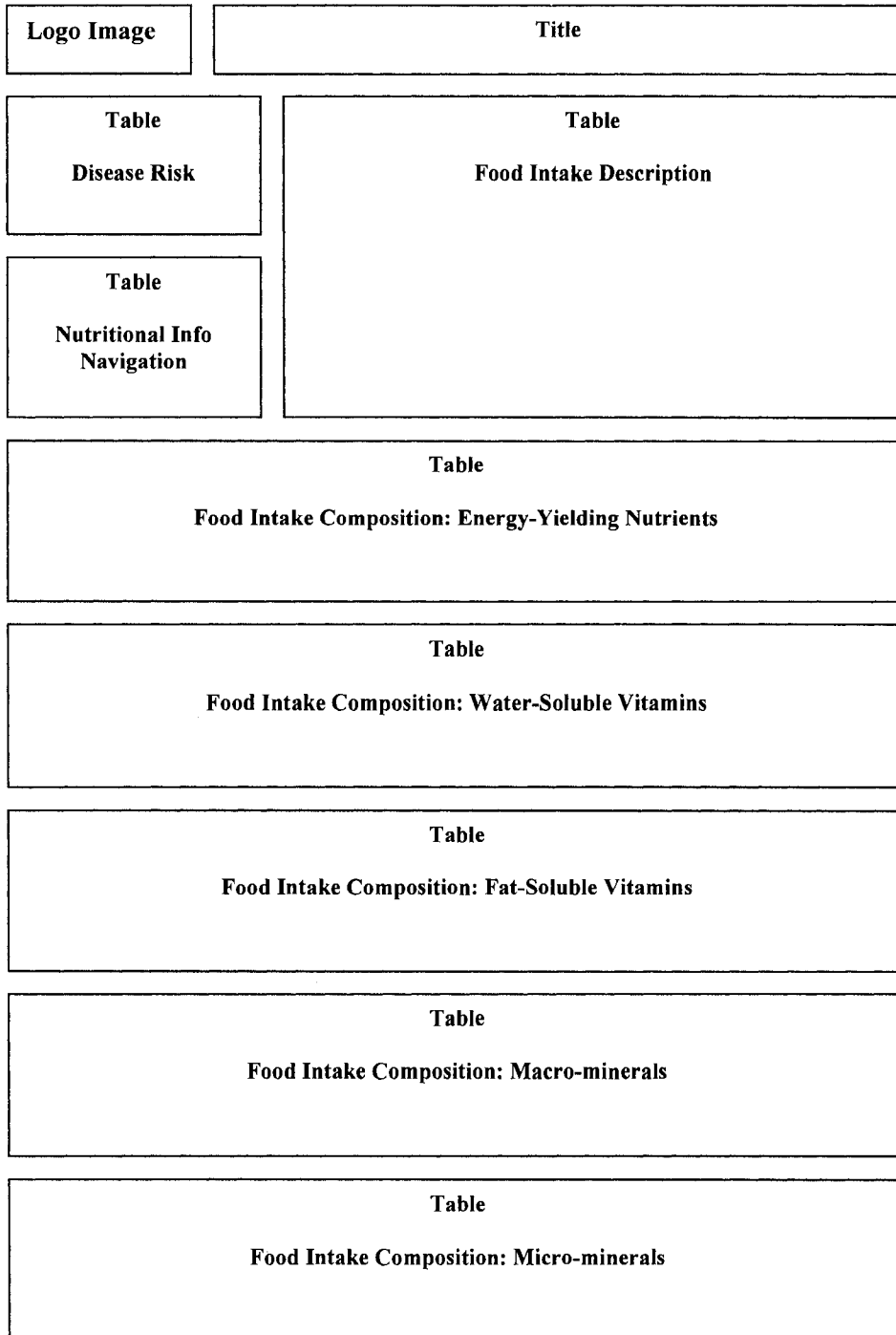


Figure 30: Food Intake Pattern XML File Layout.

their amounts and association with a particular food group. The “Disease Risk” table displays information about diet-related risk for three cardiovascular diseases: (a) heart disease, (b) stroke, and (c) hypertension. The detailed nutritional information is provided in 5 tables (1) Energy-Yielding Nutrients, (2) Water-Soluble Vitamins, (3) Fat-Soluble Vitamins, (4) Macrominerals, and (5) Microminerals. In particular, the table “Energy-Yielding Nutrients” describes each food item in the patterns in terms of its calorie, fat, saturated fat, trans fat, carbohydrate and fiber intake. Similarly, other tables provide information about food item composition in terms of vitamins and minerals. Since these tables are pretty big and may occupy a whole page in the browser, the table “Nutritional Info Navigation” provides quick links to the nutritional information tables at the top of the page. Besides that, beneath each table there is a link to return to the top of the page, if necessary.

Since food intake data are kept in a compound database with two components represented by a relational database and an XML database of pattern files, care should be taken to keep them synchronized so that there would be always a pattern file in the XML database with the file path stored in the “Pattern” table. For example, the consistency across the database components can be maintained with a consistency checker script which would detect any pattern file whose metadata are missing in the “Pattern” table data and any “Pattern” table tuple pointing to a non-existent pattern file.

2.7 System Architecture

The Reverse Risk Assessment software is supposed to be implemented as a web-based application characterized by typical 3-tier architecture with the relational database server at the back end and web application server in the middle tier. The system architecture is schematically shown in Fig. 31. This architecture is characterized by a number of advantages like (a) scalability, (b) low-cost maintenance on the client side, and (c) flexibility in terms of adding new components. However, these advantages come at a certain price. First, the complexity of an application increases due to appearance of the third tier – application server. That implies, in particular, longer debugging times since each code modification requires redeployment of the application on the server-side. Secondly, the application server must be configured to communicate seamlessly with a back-end relational database system.

Clients access the application through a browser on their workstations. The application can be deployed on any J2EE-compliant application server. In our case, the Apache Tomcat application server is selected to host the application because Tomcat is renowned for its good performance benchmarks.

The Tomcat application server communicates with both XML file store and the relational database. The database store metadata about patterns, while actual pattern files are located in the XML file store. This file store, in fact, represents an XML database since all pattern food items with their nutrient composition is readily available from the respective XML content unit.

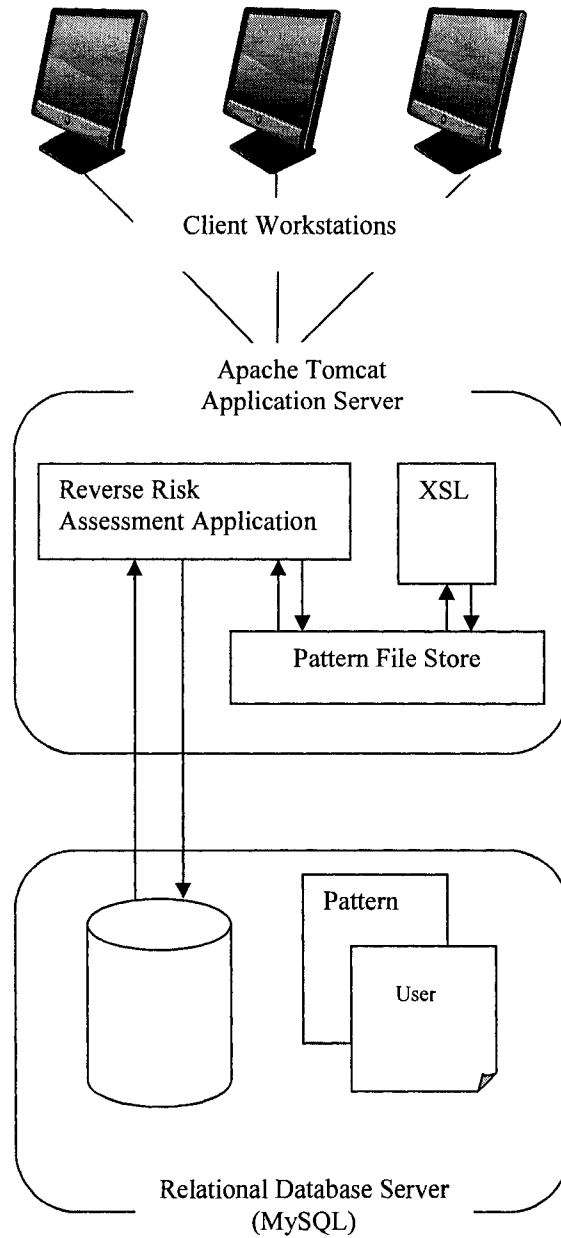


Figure 31: Reverse Risk Assessment System Architecture.

Performance consideration is another reason for using the XML database. Indeed, to perform reverse risk assessment the user runs the risk calculator which, in essence, represents a search engine. Therefore, instead of generating patterns on demand from a

food composition database, the system sends a relatively simple request to the pattern database and obtains a collection of file paths for the patterns corresponding to a selected set of risk parameters including the probability to contract a disease and diet-related risk factors. The application then provides the user with a set of patterns which satisfy his/her search criteria. The pattern set is displayed in a table which contains links to the pattern XML files. Upon clicking on the link the browser opens a separate window with a detailed food intake description and all nutritional parameters in the HTML format. The XSL transformation into the HTML format is provided by a special XSL stylesheet residing on the application server. Since opening an XML file in a browser is almost instant, the response time for risk assessment is small and the system performance doesn't suffer from the necessity to process a large number of data on demand.

2.8 Subsystem Decomposition

The system can be decomposed into four major subsystems: (1) Risk Calculator, (2) Pattern Generator, (3) User Management, and (4) General Information. These subsystems are schematically shown in Fig. 32.

2.8.1 Risk Calculator

The "Risk Calculator" subsystem provides the user with the ability to specify (a) a level of acceptable risk, (b) a disease, (c) food preferences and non-diet related risk factors.

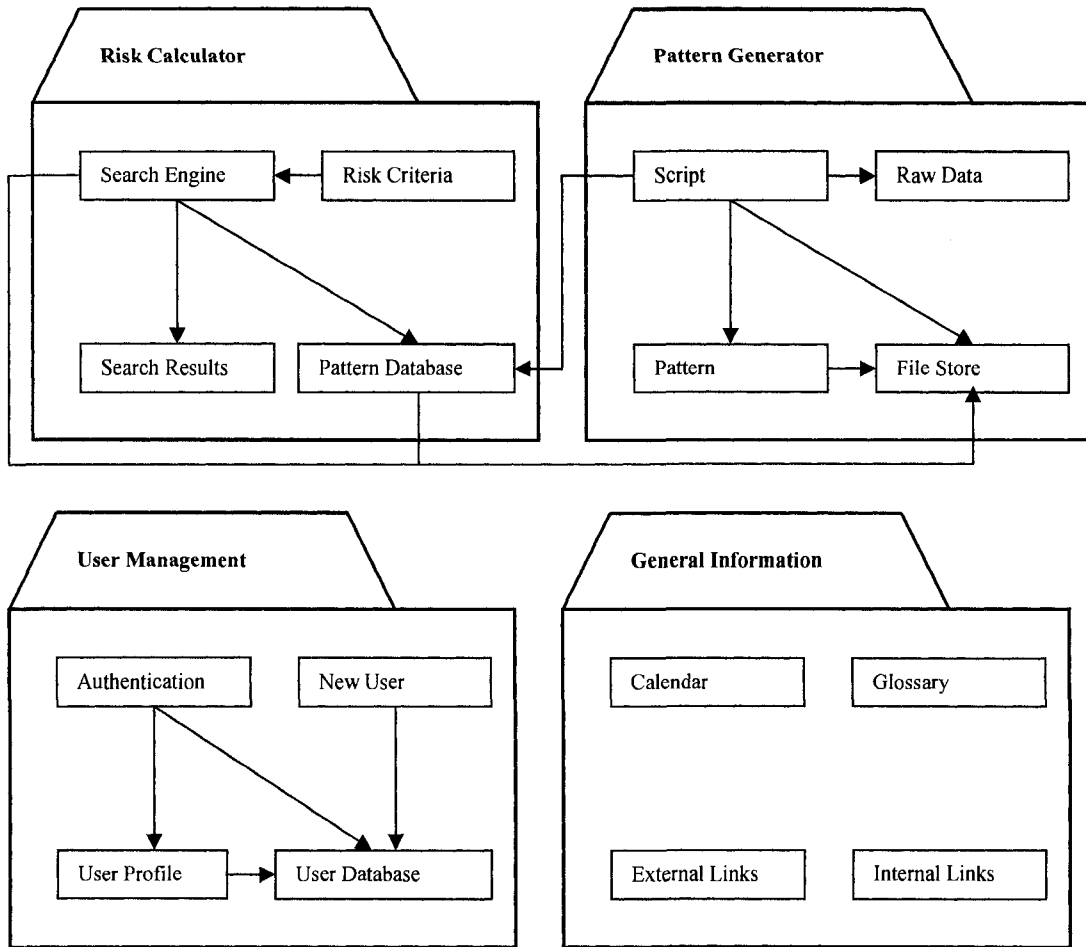


Figure 32: Reverse Risk Assessment Subsystem Decomposition.

This component also accepts a user request to generate an appropriate food intake pattern set. Therefore, logically it consists of an interface for specify pattern search criteria, a search engine which translates the criteria into a query, a relational database and an interface to display search results.

2.8.2 Pattern Generator

The Pattern Generator subsystem populates the relational database with food intake pattern metadata and the XML database with food intake pattern files in XML format. The metadata are stored in the table “Pattern” whereas pattern files are distributed across the system of folders in the XML database. The subsystem consists of four major components: (a) Raw Data, (b) Script, (c) Pattern, and (d) File Store.

The “Raw Data” component represents a set of files containing the USDA National Nutrient Database [56]. The “Script” component is a program which runs against the “Raw Data” database and generates pattern files by filling in appropriate attribute values. A high-level flowchart for the pattern generation algorithm is shown in Fig. 33. The algorithm is based on random selection of food items for each food group. The selection of food items is made from the Food and Drug Administration Database [56] which contains all relevant nutrient information. It should be emphasized that the node “Get nutrient information” in Fig. 33 implies a double loop over nutrient groups and nutrient items in each nutrient group. These loops are not shown in Fig. 33 for the sake of clarity. The risk of cardiovascular disease and other related diseases (heart disease, stroke, and hypertension) is determined by the formulas derived in the mathematical formalism developed earlier. The mathematical model calculates the probability of contracting a disease in the linear approximation based on probability gradients derived from formula (14).

When generating patterns the program employs an XML template, which is a part of the “Pattern” component. The latter also includes an XSL stylesheet, a CSS stylesheet and a DTD (Document Type Definition) file. The generated food intake patterns are

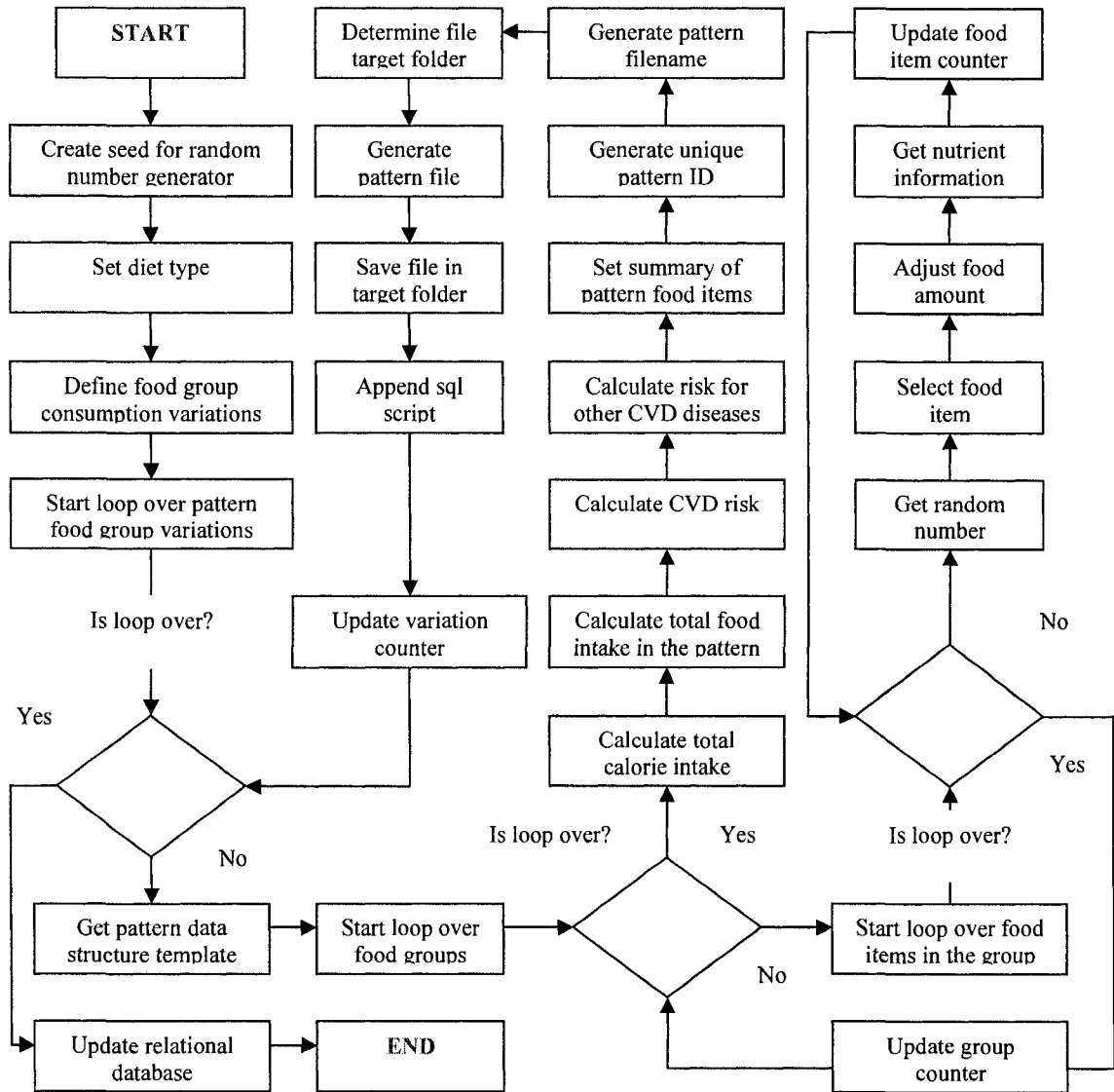


Figure 33: Food Intake Pattern Generation Flowchart.

stored in the “File Store” component which represents an XML database. Each time a pattern is generated and a respective XML is created in the File Store, the corresponding pattern metadata record is made to the relational database.

2.8.3 General Information

The “General Information” subsystem is intended for providing basic information about diet-related risk factors and chronic diseases. It also contains external links to web sites which can be useful in overall diet evaluation and risk assessment. Thus, this subsystem consists of the following components: (a) “Internal Links”, (b) “Glossary”, (c) “Calendar”, and (d) “External Links”.

The “Internal Links” component is a collection of web pages with the information about risk factors and chronic diseases. In particular, the risk-factor pages include (1) “Cholesterol”, (2) “Fat”, (3) “Alcohol”, (4) “Salt”, (5) “Sugar”, (6) “Fiber”, (7) “Vitamins” and (8) “Minerals”. Correspondingly, the disease pages describe (1) “Heart Disease”, (2) “Stroke”, (3) “Hypertension”, (4) “Obesity”, (5) “Osteoporosis”, (6) “Diabetes”, (7) “Diverticulosis” and (8) “Dental Caries”.

The “Glossary” subcomponent represents a page with a glossary of diet and nutrition-related terms with easy alphabetical navigation. The “Calendar” component consists of a page with a calendar applet and a table with the information about incoming conferences on health care, nutrition and risk assessment. Finally, the “External Links” component provides links to the external web sites with the additional information on healthy diet and lifestyle.

2.8.4 User Management

The “User Management” subsystem serves three purposes. First of all, it authenticates users, secondly, it allows ordinary users to update their own profiles once they are logged into the system, and, finally, this component also provides a system administrator with the ability to create a new user profile. Therefore, it is stipulated that the application provides interfaces for logging into the system, updating and creating a user profile. It is assumed that the system maintains a table of users authorized to access the application.

2.9 Design Class Diagrams

Since the software is conceived as a web-based application leveraging the JavaServer Faces technology (See Chapter 3), it is logical to use the extended UML model for web application design [57, 58]. It should be noted, however, that the JavaServer Faces technology imposes additional constraints on the web application profile extension to UML. First of all, each page in the application contains a form element, even though the page itself may not contain any dynamic HTML content. Secondly, there is a unique correspondence between a page and the respective Java bean which supports and encapsulates all dynamic HTML functionality of the page widgets, including simple link navigation. Thus, each JSF page form contains input elements which are represented as objects in the underlying Java bean. Therefore, the uniqueness of a page in terms of functionality is primarily determined by the page form structure and its object

representation in the Java bean. That means the total number of design class diagrams can be identified as a total number of functionally unique input forms in the application. In our case, there are 5 such forms (with self-descriptive notation): (1) “Search”, (2) “Login”, (3) “Update Profile”, (4) “New User”, and (5) “Simple Navigation”.

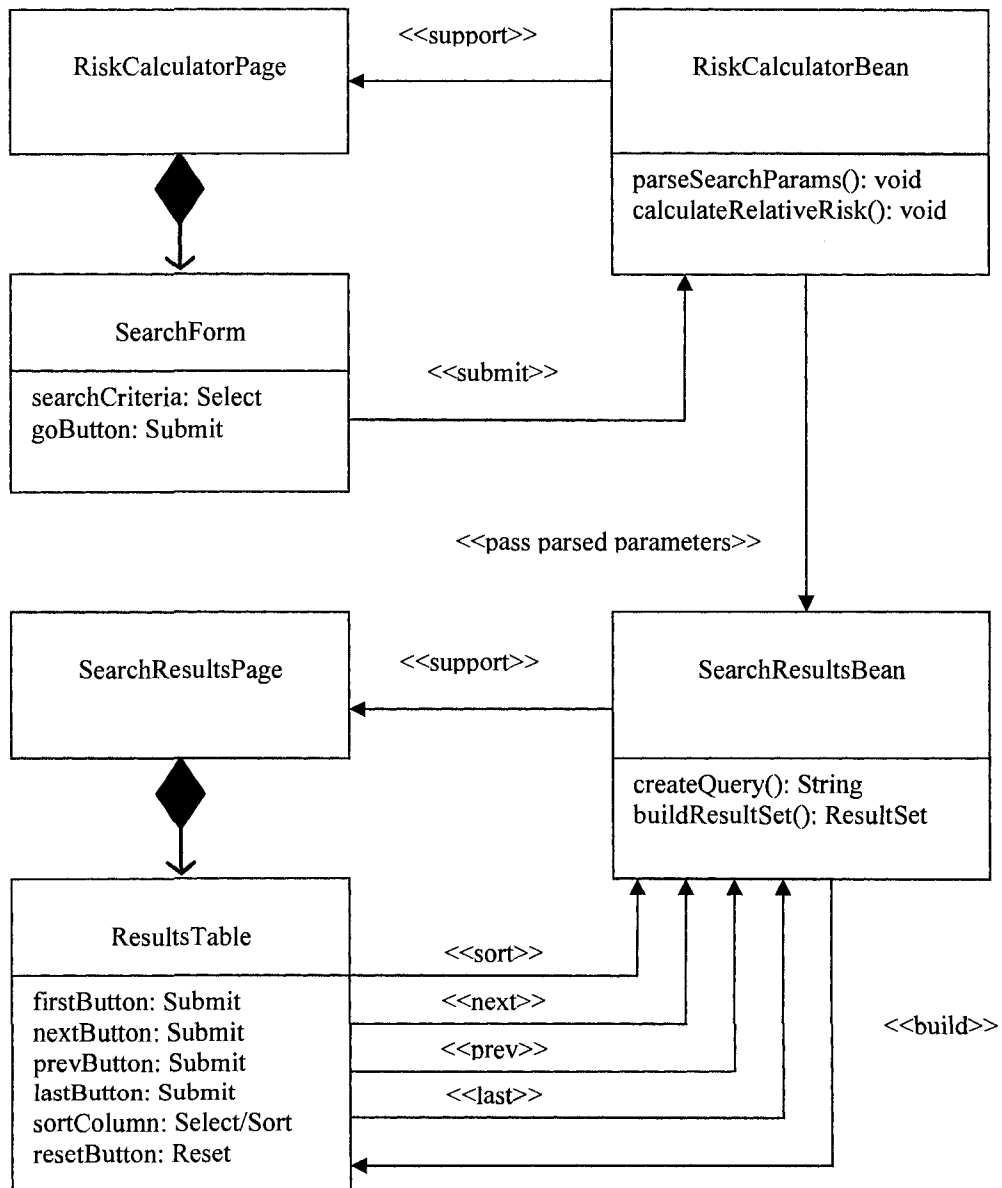


Figure 34: Risk Calculator Design Class Diagram.

The “Search” form belongs to the “Risk Calculator” page and is designated for the user input related to pattern search and risk evaluation. This page is closely related to the “Search Results” page which displays pattern result sets. The results are organized in a table which allows the user to browse the result set page by page. Besides that, the table

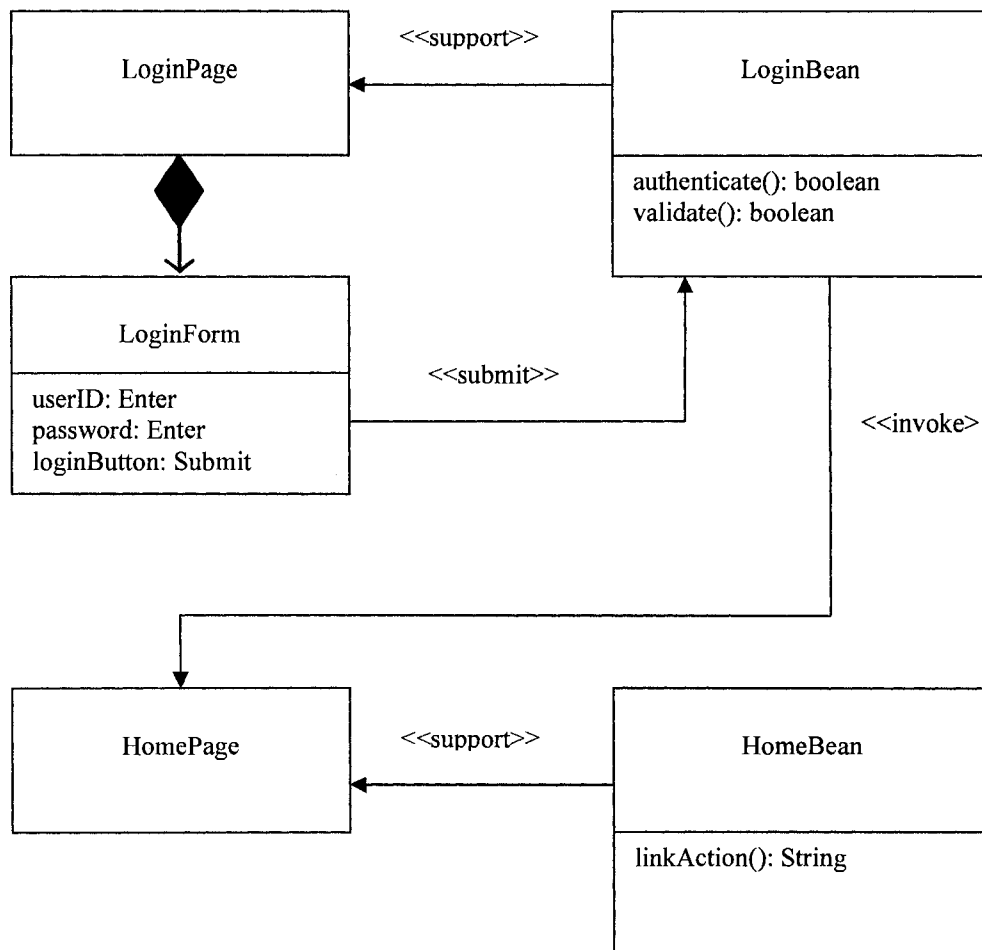


Figure 35: Login Design Class Diagram.

also provides the possibility to sort results by column or a set of columns. Figures 34 illustrate the respective design class diagram.

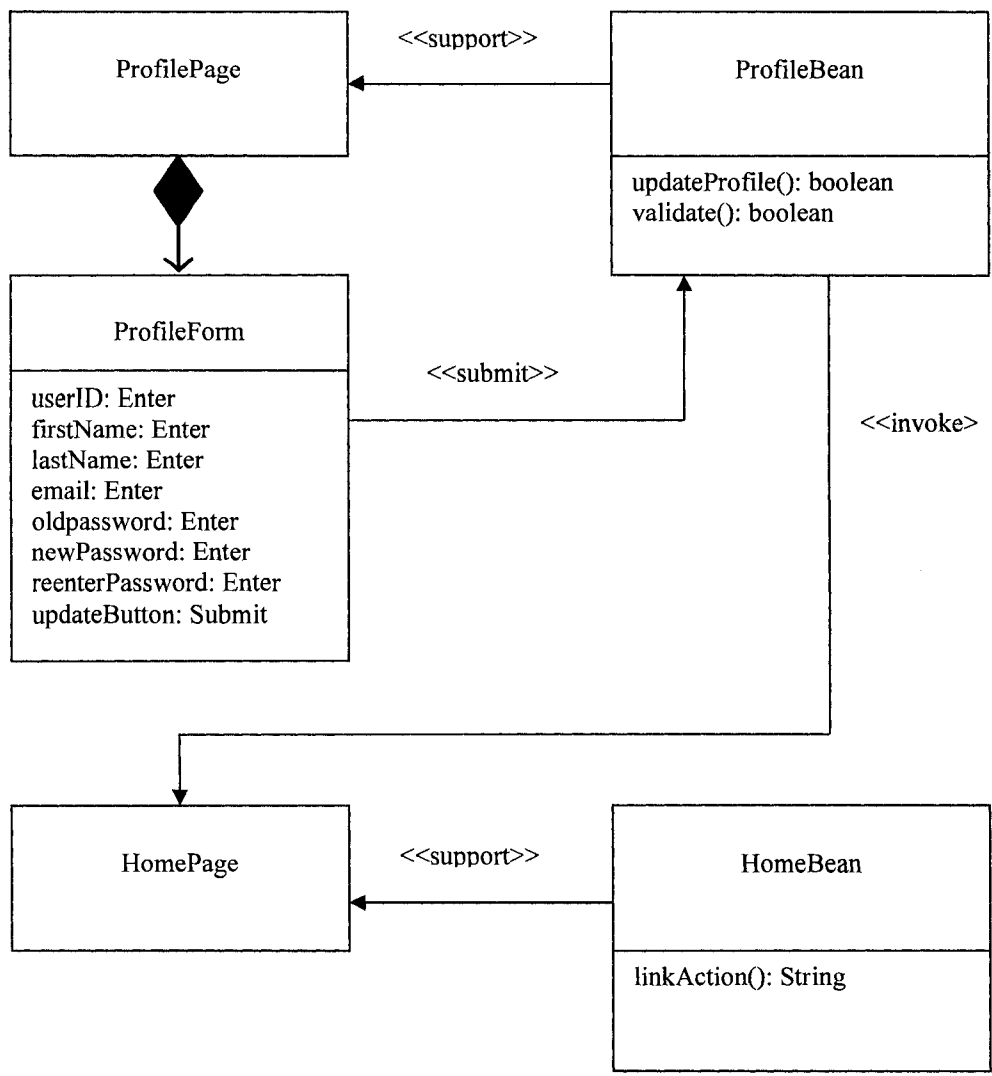


Figure 36: Update Profile Design Class Diagram.

The purpose of the “Login” page and the associated form is to authenticate the user and grant him/her access to the “Home” page, which cannot be invoked without providing valid authentication credentials. The Login design class diagram is shown in Fig. 35.

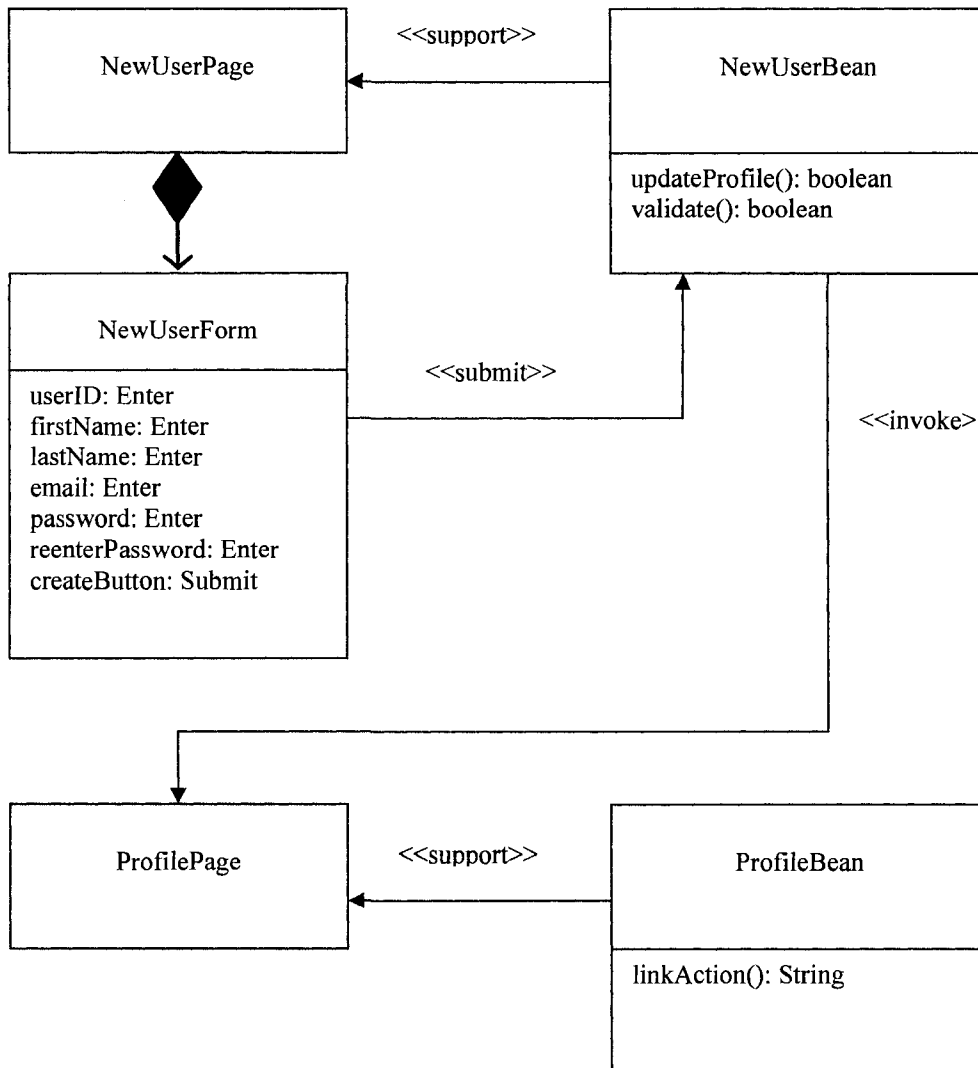


Figure 37: New User Design Class Diagram.

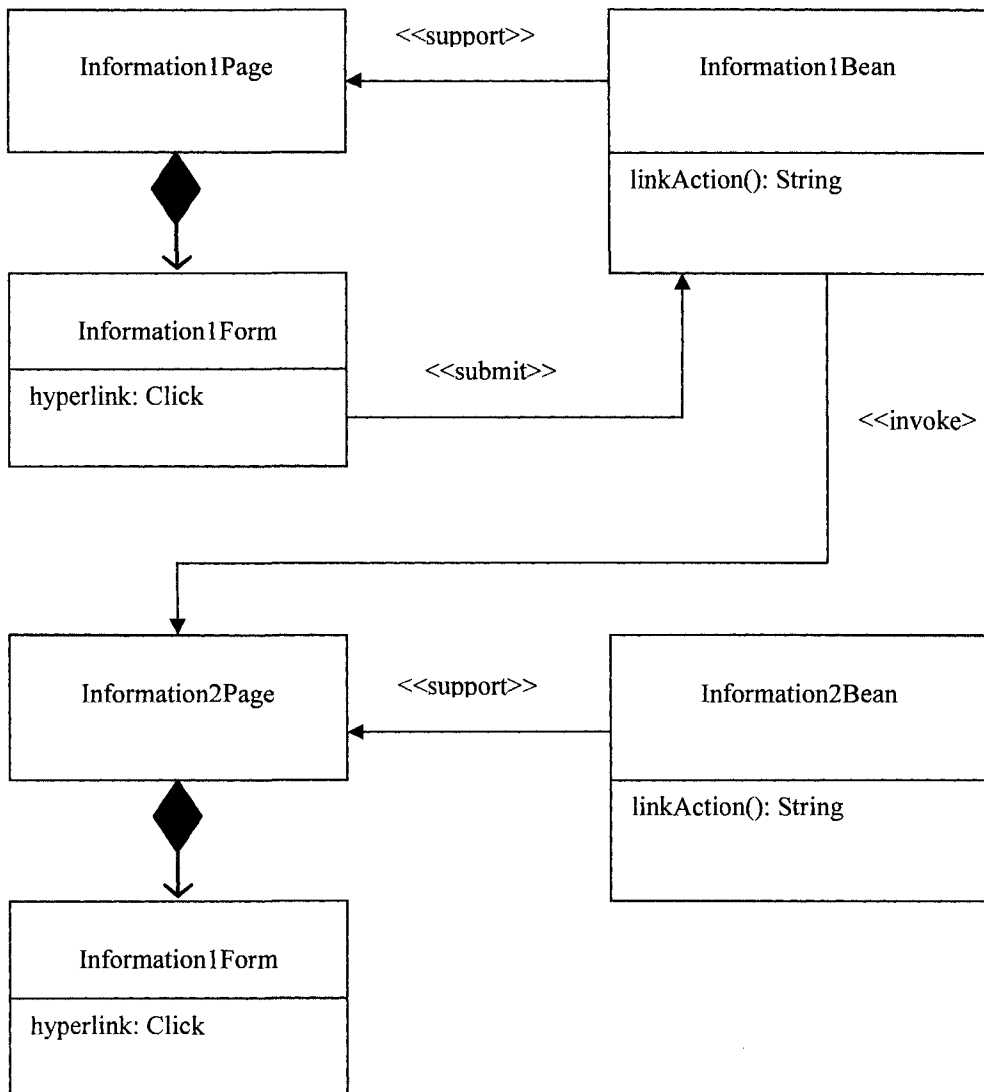


Figure 38: Simple Navigation Design Class Diagram.

The “Update Profile” page can be accessed by any user from the “Home” page and serves for updating a user profile. The corresponding class diagram is schematically illustrated by Fig. 36. The “New User” page is only accessible by a system administrator,

who must be able to create a new user profile. The corresponding design class diagram is shown in Fig. 37.

All the rest of the pages have simple forms which contain one or more instances of the only one dynamic component – “Link Action”. This component provides page navigation for properly authenticated users and returns a string parameter pertaining to the next invoked page. Fig. 38 shows a design class diagram for handling simple navigation across informational pages of the application. For the sake of generality, the generic notation is used to denote two informational pages connected through a hyperlink action.

3. Implementation and Risk Assessment Results

This chapter describes the reverse risk assessment system implementation and risk assessment results obtained by generating a large set of food intake patterns.

3.1 Java Server Faces Technology Overview

The Java Server Faces technology has been selected as a framework for implementing the reverse risk assessment application. There are several reasons for this choice. First of all, JSF simplifies the design and implementation of GUI by offering a rich set of reusable UI components. Secondly, JSF provides transparent linking of those UI widgets to server-side event handlers and data sources [59, 60]. Thirdly, this technology ties closely together web UI programming and Java Beans. Another important consideration is that JSF is an open source technology which evolved by taking the best ideas and practices from the already existing web development techniques [61, 62].

The JSF framework supports the traditional Model-View-Controller architecture which allows web page designers to create a user interface, on one hand, and application developers to focus on the application logic, on the other. However, the power of JSF is in its ability to directly link GUI elements with server-side code. JSF literally “animates” the old static HTML widgets and makes them extremely versatile. JSF also provides

services for data conversion, validation and error handling, which significantly reduce development time.

The most recent predecessor of JSF is the Struts technology which has already reached its stage of maturity. The JSF technology has a number of advantages over Struts which are sometimes are not clearly understood. First, as opposed to JSF, there is only one event handler per each HTTP request in Struts. As soon as the request is satisfied and the Struts command controller did its job, a response is returned to the Struts front controller which, in turn, decides on the next navigation step. In JSF, several events can be handled in one request since JSF utilizes a page controller paradigm [63]. Secondly, the Struts command controller uses the extended Action classes to process the request which translates into a Struts ActionForm. Extension of Action classes to support ActionForm handling adds an additional layer of complexity and is often a source of bad coding practice. In JSF, a developer just has to customize an automatically generated Java Bean which supports the respective JSF page. Thirdly, Struts has been oriented from the beginning at the JSP/HTML approach (which was logical at the time the Struts technology was born) and has not been intended to support a fully developed standard component model. The JSF framework, the UI components are made readily available to developers and significantly facilitate the RAD process.

There are two notions which make JSF distinctly different from all the predecessor technologies. Those are the component tree and the JSF lifecycle. They are briefly discussed in the next two sections, respectively.

3.1.1 JSF Component Tree

One of the JSF concepts which play an important role in rendering a page to the user is the component tree. The component tree represents a data structure which combines all GUI widgets on a JSF page [60]. The widgets are defined and described by JSF tags which are pretty similar to JSP tags. However, the difference between JSP and JSF lies in many respects in the component tree. In JSP all tags are processed sequentially, from the beginning to the end, without any noticeable dependence between each other. Correspondingly, JSP actions are processed in the order the respective tags are read. On the contrary, the tags on a JSF page are arranged in a hierarchical order and each of them has a built-in handler class. In this respect, a JSF page may be considered as a visual and run-time representation of a component tree. Each time a browser sends a request to the server, the JSF servlet initializes the page component tree code to deliver the page in the browser as if it were a normal HTML page.

3.1.2 JSF Lifecycle

In general, a JSF application is supposed to pass the client's HTTP request to the server, process the request on the server side and render the response. In more detail, the JSF lifecycle consists of six major phases which include (1) Restore View, (2) Apply Request Values, (3) Process Validations, (4) Update Model Values, (5) Invoke Application, and (6) Render Response [60].

In the "Restore View" phase the request is received and the component tree is reconstructed if the respective page was displayed for the first time or retrieved if the

page was displayed previously. In case the request carries no data, phases 2-5 are skipped and the response is delivered in phase 6. An example of this action would be a simple navigation using HTML links.

If the request contains data the lifecycle proceeds to the next stage which is “Apply Request Values”. During this stage, the component tree objects are scanned for the request values which are stored for further processing. The first step in request data processing consists in converting submitted string values to “local” values of appropriate types and validating those values. If there are any validation errors, further processing is aborted and the lifecycle goes to stage 6 – “Render Response”. When there are no validation errors, the lifecycle proceeds to the stage “Update Model Values”. The latter is used to update the Java Beans linked to the components. After component’s objects are updated, the “Invoke Application” phase takes place and actions associated with buttons and links are executed. Finally, the J2EE web container traverse the component tree and renders an appropriate page in the Render Response” phase.

Technically, all request management and lifecycle processing in JSF applications is handled by the FacesServlet which is not directly accessible to developers. The FacesServlet stores information necessary for lifecycle and request processing in the object called FacesContext [64]. It should be stressed that although the FacesServlet supports two types of requests and responses: (a) native JSF and (b) non-JSF requests/responses, their relationships with the JSF lifecycle are different. . The native JSF requests originate from a JSF page and the respective JSF responses are rendered during the “Render Response” phase. The non-JSF requests are sent to application

components which are not a part of the JSF framework. Correspondingly, a non-JSF response is not rendered by the “Render Response” phase [64].

As one can imagine, JSF development may be inefficient without an appropriate Integrated Development Environment (IDE). This IDE is discussed in the next section.

3.1.3 Java Studio Creator 2

Using JavaServer Faces framework with traditional development tools like Eclipse, JDeveloper, JBuilder or NetBeans may be difficult and inefficient since those tools have not been designed to work directly with JSF web-based applications. Therefore, Sun Java Studio Creator IDE which was specifically implemented and tailored by Sun Microsystems for JSF developers. The first version of this IDE named Sun Java Studio Creator (JSC) was made available to the users at the beginning of 2004 and is known as Creator 1 [65]. Few years later, a significantly enhanced version of the product, Sun Java Studio Creator 2, was released to the market. This latest version of Creator proved to be much more powerful and fixed a lot of bugs present in the original release.

The Java Studio Creator IDE is an ideal choice for medium-size applications like the one developed in this project. The following features, in particular, have been taken into account while selecting this tool:

- (1) A palette of standard GUI components with the drag-and-drop capability.
- (2) Convenient navigation utilities for accessing various application components.
- (3) A utility for fast database connection configuration with a number of out-of-the-box drivers from leading database management system vendors.

- (4) A database view which allows the user to execute SQL queries directly from the IDE. This feature is extremely useful when debugging and testing database response.
- (5) Direct application deployment on the web application server. The tool comes with built-in Sun Application Server which can be used for quick deployment and debugging. Even though restarting the server takes time, the mere fact that all the rest of the work like compiling, packaging, and deploying a war file can be done in one click saves a lot of effort.
- (6) An option of importing existing HTML pages into a JSF application. This option provides a great deal of flexibility for developers since the JSC can be used in combination with the best-of-breed HTML editors like Dreamweaver, Contribute and others for the initial web page design.
- (7) An intelligent code editor with synchronized editing and debugging. The editor provides key word colorization, handy access to JavaDocs entries and automatic expansion/collapse of code segments.

Apparent simplicity of usage of the JSC IDE may create a deception impression that the tool is designed for developers with little programming experience. In reality, the JSC offers a rich palette of options which require a deep knowledge of a broad range of technologies like J2EE, SQL, JavaScript, AJAX, Web Services to name a few. In particular, the web-applications developed with the JSC tool can natively utilize the power of Enterprise Java Beans (EJB). On another note, the SQL query editor enables the user to build queries visually by dragging and dropping tables into the editor window.

Another important consideration is that JSC generates JSF pages in pairs with the supporting Java beans and keeps a map of all GUI objects behind the scenes. Therefore, manual editing of JSF pages must be done carefully with a good understanding of the underlying technology. For example, when importing existing HTML pages and blending them into an application, care should be taken to follow a strictly defined procedure. Otherwise, it is easy to run into problems and get caught by unintentional code changes.

3.2 Apache Tomcat Application Server

Being composed of servlets, JavaServer Faces, Java beans and event listeners, the reverse risk assessment application is intended to run on a J2EE web server or any application server which includes a web container. In this respect, Apache Tomcat application server [66] developed as a reference implementation of JavaServlet and JSP specifications seems to be a logical choice for an application deployment platform. Indeed, the reverse risk assessment application doesn't require complex business and data tiers of a full-fledged application server. Therefore, the runtime environment with naming context and lifecycle management services provided by Tomcat is quite adequate for running the application. The Apache Tomcat software is an open source code which is readily available for downloading from the Apache Foundation web site.

From the application deployment and system setup point of view, there are three Tomcat-specific features which needs to be addressed here: (a) installation requirements,

(b) class loading, and (c) database connectivity. They are briefly discussed in the sections below.

3.2.1 Tomcat Installation and Directory Structure

Installation of the Tomcat requires that an appropriate Java Runtime Environment (JRE) be installed first on the application server host. The Tomcat installer should also match the operating system since the Tomcat installer is OS-dependent. In this project, Tomcat is installed on Windows OS in the default directory “C:\Program Files\Apache Software Foundation\Tomcat 5.0”. This directory is commonly referred to as Tomcat home directory and the CATALINA_HOME environmental variable is used to point to this folder. The directory structure beneath the Tomcat home directory include the nine folders (a) bin, (b) common, (c) config, (d) logs, (e) server, (f) shared, (g) temp, (h) webapps and (i) work. The “bin” folder contains major executables, batch and shell script files, whereas the “webapps” directory is a folder designated for web application deployment. When Tomcat finds a new web archive file in this directory, the archive file is automatically expanded into the appropriate directory hierarchy beneath the “webapps” folder. The “config” folder is used for system configuration files while the “logs” folder stores log information. The “common” and “shared” folders are intended for installing shared libraries and are employed by the class loader utility. The process of class loading plays a crucial role in any application execution. Therefore, the next section discusses some of the most important characteristics of Tomcat class loading.

3.2.2 Class Loading

Upon initialization, Tomcat creates 6 class loaders: (1) Bootstrap, (2) System, (3) Common, (4) Catalina, (5) Shared, and (6) WebappX [66]. The Bootstrap loader is the first class loader to be invoked. It loads all the necessary classes provided by Java Virtual Machine and that is one of the reasons why JRE must be installed prior to Tomcat installation. The System class loader was intended initially for loading the class libraries from the CLASSPATH environmental variable. However, most current implementations of Tomcat use the start-up scripts which ignore the system CLASSPATH value. Instead the scripts load the classes from Tomcat internal libraries and the libraries from the tools.jar file which is a part of JRE installation. Therefore, a web application should not rely on the CLASSPATH value.

To bridge the gap and add more flexibility in terms of common libraries access, the Common class loader makes classes and libraries located in the “common” folder visible to both Tomcat executables and web applications. Classes which are exclusively used by Tomcat are loaded by the Catalina class loader. Those classes are protected and cannot be accessed by any web application. The Shared loader is somewhat similar to Common except for it loads classes from the “shared” folder. It is assumed, however, that no Tomcat classes are stored in the “shared” folder. Finally, the WebappX class loader initializes a separate loader instance for each application deployed on Tomcat. The respective loader instance loads the classes from the WEB-INF/lib subdirectory beneath the web application root directory.

3.2.3 Database Connectivity

The configuration procedure which ensures proper communication between Tomcat and a relational database server comprises of two steps. First, it is necessary to place an appropriate driver jar file into the “common/lib” subfolder. Secondly the Tomcat configuration file “server.xml” located in the “config” folder must be modified to include an entry for the corresponding data source. The details related to the specific database driver implementation and the server.xml file modification are described in the section “System Specifications and Deployment”.

3.3 MySQL Relational Database Management System

MySQL relational database management system has been selected as a backend database for the reverse risk assessment system. Although MySQL is an open source software, it is a robust SQL database server capable of handling efficiently multiple user connections [67]. In particular, the following considerations have been taken into account when selecting MySQL as a backend database system:

- (1) Performance. Written in C and C++, MySQL uses fast and highly optimized methods of data retrieval, as well as thread-based memory allocation. The database response for simple queries is within the range of few seconds or even better.

- (2) Portability. MySQL can be deployed on many operating system including Windows, Unix and Linux.
- (3) Scalability. The database server is able to handle very large databases with up tens of millions of records.
- (4) Security. The system offers a flexible password-based system of privileges for database access and configuration. Upon transmission across the network, passwords are encrypted.
- (5) Data Types. The database system supports all major data types.
- (6) SQL Statement Support. The system supports all standard SQL statements including “select”, “delete”, “insert”, “replace”, and “update”. It also provides support for aggregate functions like count(), count(distinct...), avg(), sum(), etc.
- (7) Database Connectivity. MySQL clients can be written in a variety of languages such as C Sharp, C++, Java, et al. In particular MySQL JDBC driver is available as open source software.
- (8) MySQL Command Line Console. This utility allows one to quickly access the database objects from the command line interface. It is also possible to execute scripts in a batch mode.

The MySQL benefits listed above made it popular in many three-tier systems utilizing Linux, Mac or Windows OS in combination with Apache web server or Apache Tomcat. It is estimated that currently there are 10 million MySQL installations. Recently it was

announced that MySQL AB, the company which produces MySQL, would be acquired by Sun Microsystems by the end of June 2008.

3.4 Application Directory Structure

The Reverse Risk Assessment application directory structure follows the rules for a J2EE web application archive. Thus, all JSP pages are placed immediately below the application root folder called “ReverseRiskAssessmentApplication”. Figure 39 shows the application folder hierarchy and some of the JSP pages in the right-hand side pane. As already mentioned, each JavaServer page is supported by a respective Java bean. The Java bean classes are packed into the “reverseriskassessment” package which is placed into the “classes” subfolder beneath the WEB-INF folder. Fig. 40 shows the application package class files in the alphabetical order. Running a JSF application always requires a number of external libraries. Those libraries can be placed into several folders on Tomcat as follows from discussion in the section “Tomcat Installation and Directory Structure”. However, storing the libraries in the “lib” subfolder beneath the “WEB-INF” folder, as shown in Fig. 41, has a number of advantages. First of all, the application web archive becomes completely portable and no library-specific configuration is required to run the application once the war file is copied to the “webapps” directory. Secondly, all the

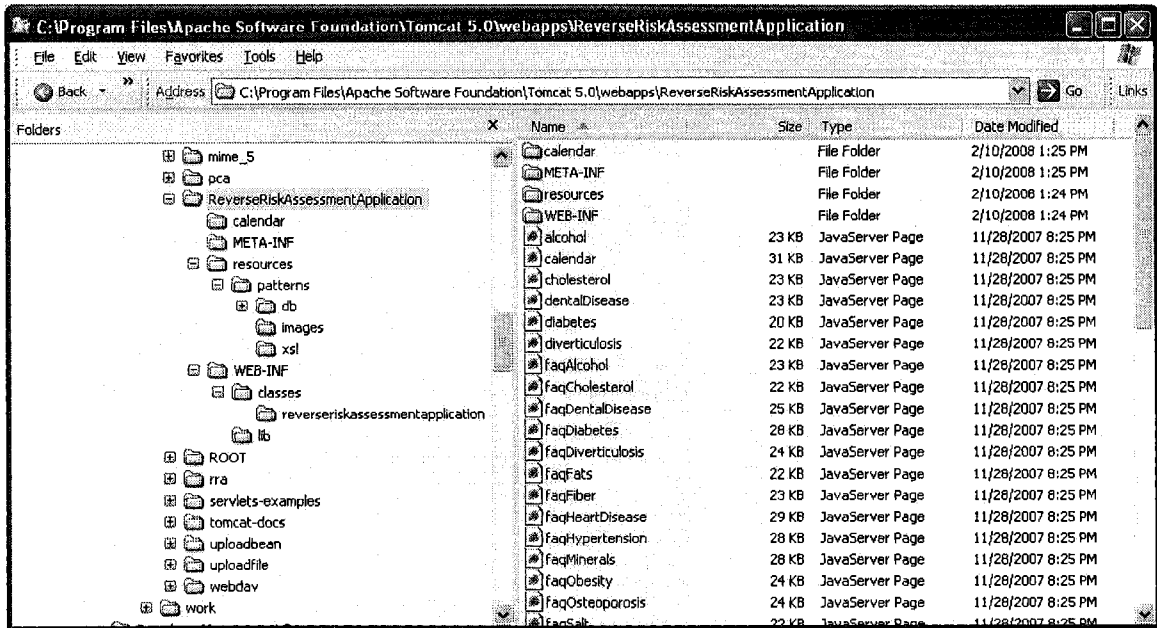


Figure 39: Application Folder Hierarchy.

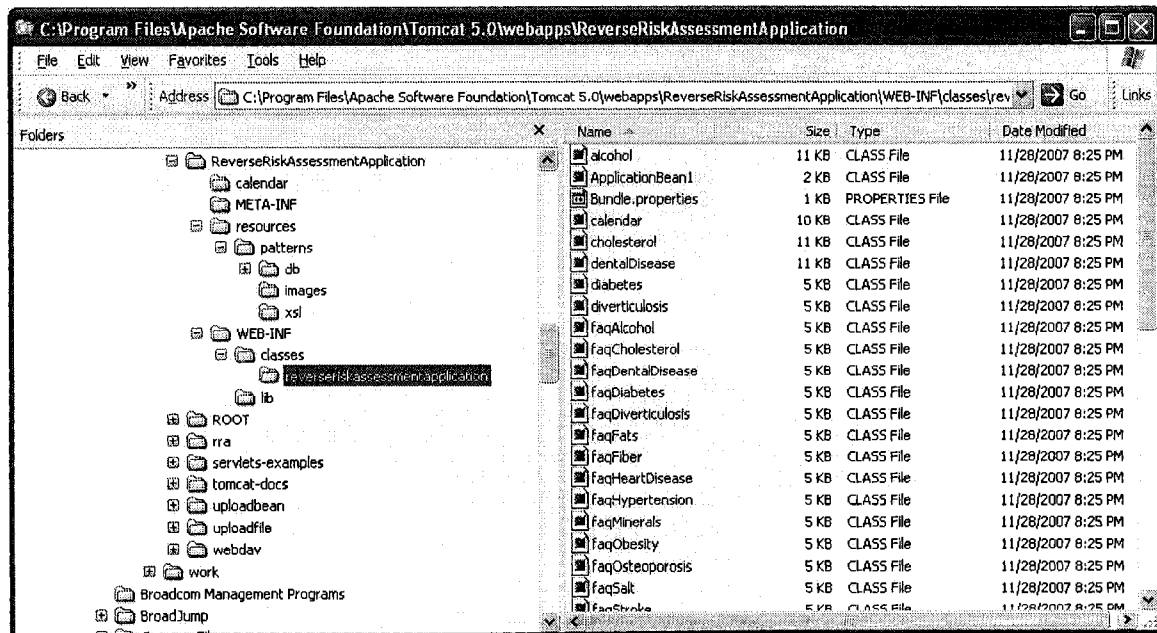


Figure 40: The Application Package Class Files.

libraries pertaining to the JSF application, including the driver jar file “mysql-connector-java-3.0.15-ga-bin.jar”, are clearly visible and localized in one folder. Note that Java Studio Creator automatically places a number of important JSF libraries like jsf-api.jar, jsf-impl.jar and jsfc.jar into the “lib” subfolder.

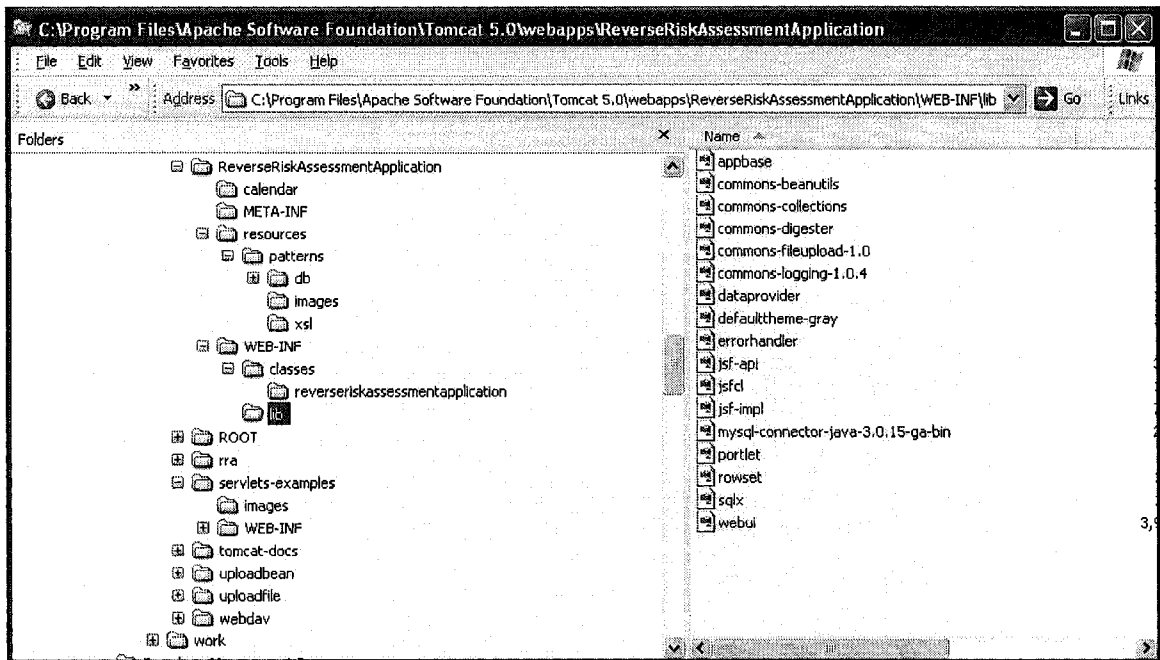


Figure 41: External Libraries Folder.

The “calendar” subfolder conveniently contains all the classes and files to support the functionality of Calendar applet, whereas the “resources” subfolder is designated for storing images. The “resources” subfolder also contains the folder “patterns” which holds resources to support graphical representation of food intake information and has three subfolders: (1) db, (2) images, and (3) xsl. The food intake pattern XML files are stored in the folder hierarchy beneath the “db” folder while the “images” folder stores images

related to the pattern representation. Finally, the “xsl” folder holds the dtd file, css and xsl stylesheet files.

3.5 Graphical User Interface Implementation

The application GUI was first designed and built in the Dreamweaver HTML editor. Then the HTML files were imported into the application project using the built-in utility in Java Studio Creator IDE. The next section provides an overview of the application page set.

3.5.1 Web Page Set Overview

When a user invokes the application URL in the browser, the first page loaded is the “Login” page (login.jsp). The application opening page, as it appears in the browser window, is shown in Fig 42.

After being successfully authenticated the user is presented with the application Home page (index.jsp) which is illustrated by Fig. 43. The home page provides general information about the mortality rate and death statistics of the top ten causes of death in the US, and emphasizes the importance of the diet-related risk factors and healthy lifestyle. The vertical navigation on the left side of the page allows the user to connect to 16 informational pages related to risk factors and chronic diseases. As an example, Figs. 44-47 show some of the informational pages. Another vertical navigation bar, on the right-hand side of the Home page contains links to the external web sites.

Through the horizontal navigation bar, the home page also provides links to Profile, Risk Calculator, Calendar, Glossary and Login pages. The “Profile” page allows the user to update his/her profile (Fig. 48). This page also contains a link to the “New User” page, which is accessible only by the system administrator. Fig. 49 shows the “New User” page with the widgets to create a new profile on the left-hand side and the user profile look-up table which displays a listing of registered users with major user profile attributes.

The Calendar and Glossary pages are shown in Figs. 50 and 51, respectively. The Calendar page features a calendar applet and a table of incoming events and conferences on nutrition and health care. The Glossary page contains entries with explanation of different terms related to nutrition, dietary risk factors and chronic diseases. A system of alphabetical links provides convenient navigation across the glossary.

Risk assessment is performed by navigating to the “Risk Calculator” page shown in Fig. 52. A user can select a risk level in the range 0 – 100 %, one of the cardiovascular diseases and a diet type. It is also possible to narrow the search for the food intake patterns corresponding to the specified risk level by selecting values for energy, fat/saturated fat and sodium intake. In addition to that, the search criteria may include conditions on the behavioral risk factors like smoking, obesity and physical inactivity. For instance, it is possible to specify a number of cigarettes smoked per day, the body mass index (BMI) and time spent on exercising every week. Upon clicking the “Submit” button, the application retrieves the respective pattern set and displays results in the “Search Results” page. In addition to the pages shown in Figs 42-53, there are 12 other

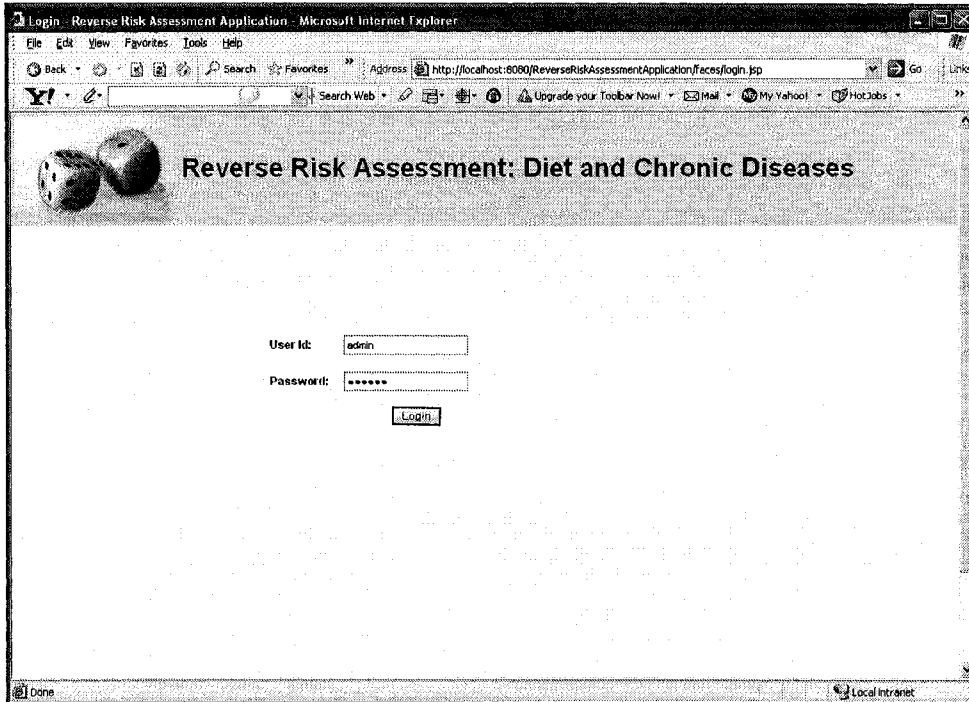


Figure 42: Login Page.

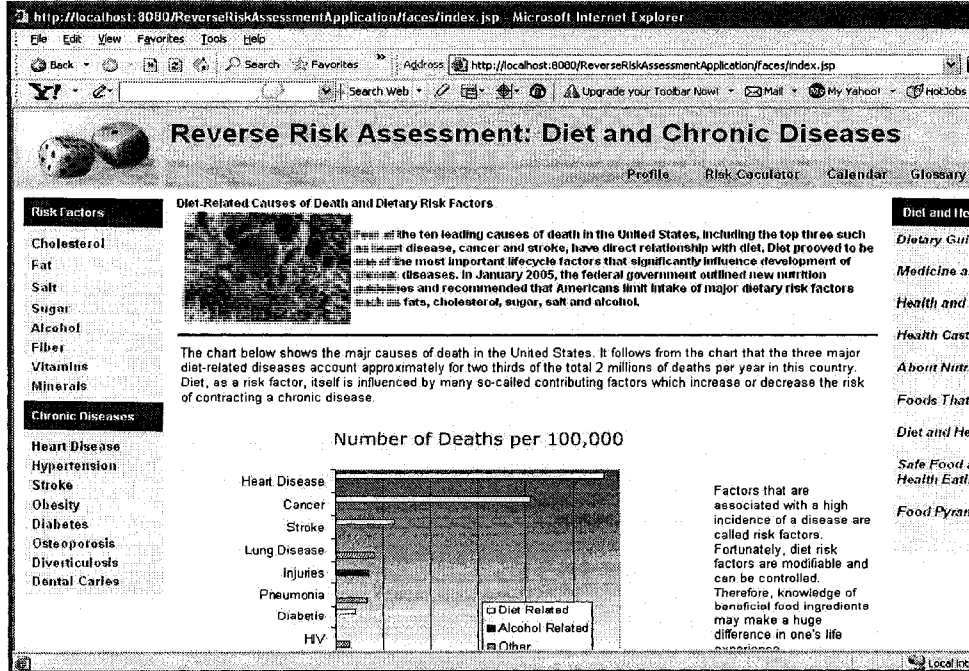


Figure 43: Home Page.

http://localhost:8080/ReverseRiskAssessmentApplication/faces/cholesterol.jsp - Microsoft Internet Explorer

Reverse Risk Assessment: Diet and Chronic Diseases

Home Risk Calculator Calendar Glossary Logout


Risk Factors

- Cholesterol
- Fat
- Salt
- Sugar
- Alcohol
- Fiber
- Vitamins
- Minerals

Chronic Diseases


- Heart Disease
- Hypertension
- Stroke
- Obesity
- Diabetes
- Osteoporosis
- Diverticulosis
- Dental Caries

Cholesterol



The gradient of cholesterol over caloric intake, which is amount of cholesterol per energy consumed, has been proved to be linearly proportional to the risk of death from heart disease and stroke [1]. Dietary cholesterol usually represents no more than 10% of all cholesterol transported by blood as most cholesterol is synthesized internally. Yet, even small increase of cholesterol concentration in blood vessels turns out to be a vital factor in developing vital chronic diseases.

Cholesterol is a soft, waxy substance that is present in all parts of the body including the nervous system, skin, muscle, liver, intestines, and heart. It is made by the body and obtained from animal products in the diet. Cholesterol is manufactured in the liver and is needed for normal body functions including the production of hormones, bile acid, and vitamin D. In particular, the cholesterol molecule is very similar to that of vitamin D3 (See picture below). As a consequence, there is no apparent need for cholesterol in the diet. There is no evidence of any beneficial effects of dietary cholesterol. Excessive cholesterol in the blood contributes to formation of deposits on the artery walls, leading to atherosclerosis and subsequent heart disease. The risk of developing heart disease or atherosclerosis increases as the level of blood cholesterol increases.



Dietary cholesterol is found in meat, poultry, seafood and dairy products. Foods from plants - such as fruits, vegetables, vegetable oils, grains, cereals, nuts and seeds - don't contain cholesterol. Egg yolks and organ meats are high in cholesterol. Shrimp and crayfish are somewhat high in cholesterol. Chicken, turkey and fish contain about the same amount of cholesterol as do lean beef, lamb and pork.

Questions and Answers

What is the relationship between cholesterol intake and health?

The relationship between cholesterol intake and LDL cholesterol concentrations is direct and progressive, increasing the risk of CHD. [More >>](#)

What is average actual intake of cholesterol in this country?

Mean cholesterol intake in The United States is above the recommended level of

Figure 44: Informational Page - Cholesterol.

http://localhost:8080/ReverseRiskAssessmentApplication/faces/fat.jsp - Microsoft Internet Explorer

Reverse Risk Assessment: Diet and Chronic Diseases

Home Risk Calculator Calendar Glossary Logout

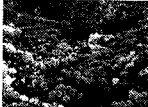
Risk Factors

- Cholesterol
- Fat
- Salt
- Sugar
- Alcohol
- Fiber
- Vitamins
- Minerals

Chronic Diseases

- Heart Disease
- Hypertension
- Stroke
- Obesity
- Diabetes
- Osteoporosis
- Diverticulosis
- Dental Caries

Fat




High fat intake is believed to be one of the major flaws of westernized diet, contributing to heart disease, obesity, development of cancer and other health problems. It has been proved that a diet low in both fat and saturated fat [1-3], combined with regular fish consumption, produces an optimal blood lipid profile [1].

Fat is a subset of the class of nutrients known as lipids. The lipid family includes triglycerides (fats and oils), phospholipids, and sterols. Triglycerides provide the body with energy and protect it from cold and mechanical shock. Phospholipids and sterols contribute mostly to cell structure and serve as a raw material for some hormones, vitamin D and bile.

A good indicator of a healthy fat intake is the blood lipid profile. A desirable blood lipid profile is characterized by four major components:

- * Total cholesterol < 200 mg/dl (5.2 mmol/L)
- * LDL cholesterol < 130 mg/dL (3.4 mmol/L)
- * HDL cholesterol > 35 mg/dL (0.9 mmol/L)
- * Triglycerides < 200 mg/dL (2.3 mmol/L)



Questions and Answers

What is the relationship between saturated fat and health?

The relationship between saturated fat intake and LDL cholesterol is direct and progressive, increasing the risk of cardiovascular disease. [More >>](#)

What is the relationship between trans fat intake and health?

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore

Figure 45: Informational Page - Fat.

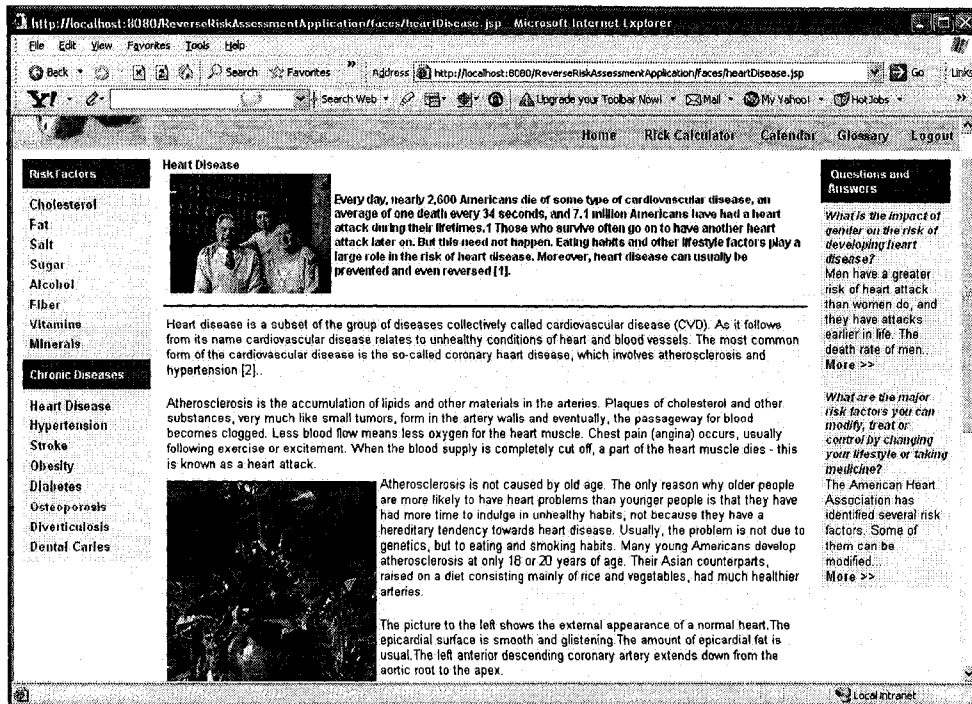


Figure 46: Informational Page - Heart Disease.

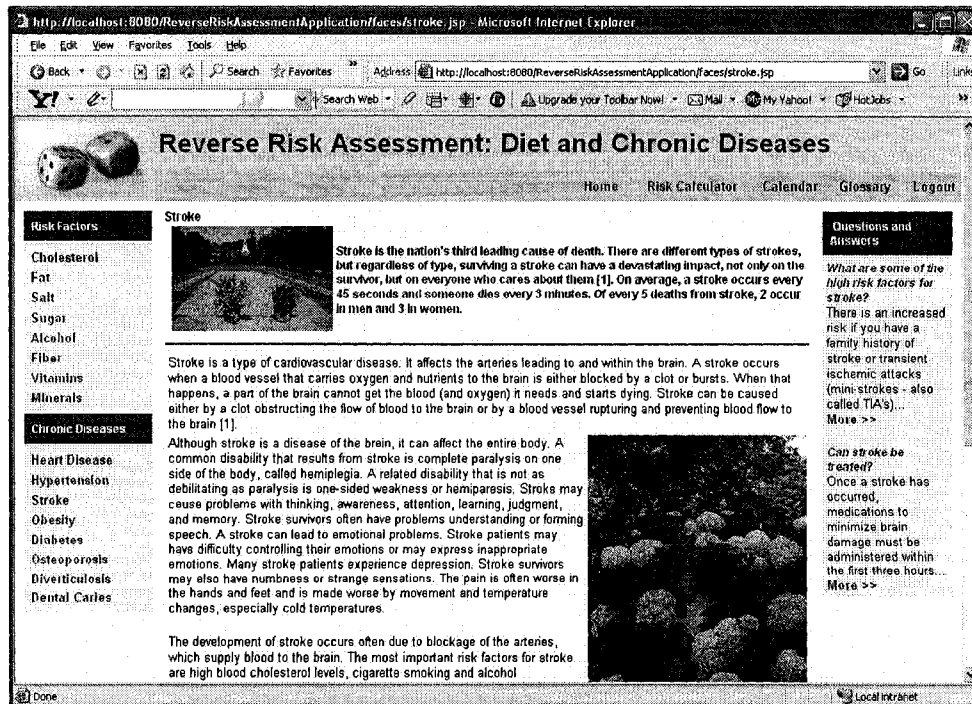


Figure 47: Informational Page - Stroke.

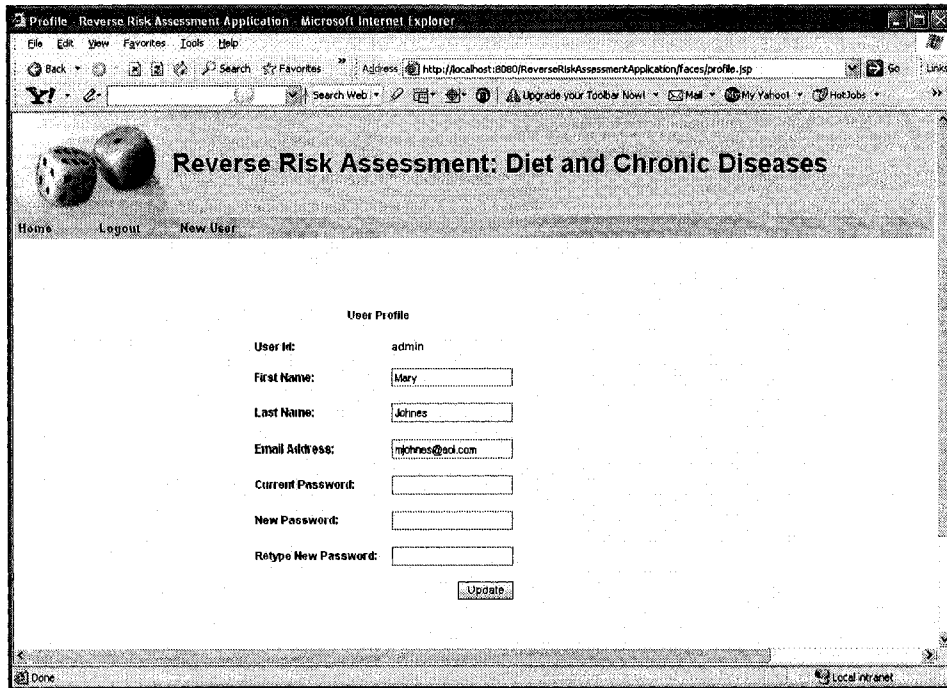


Figure 48: Profile Page.

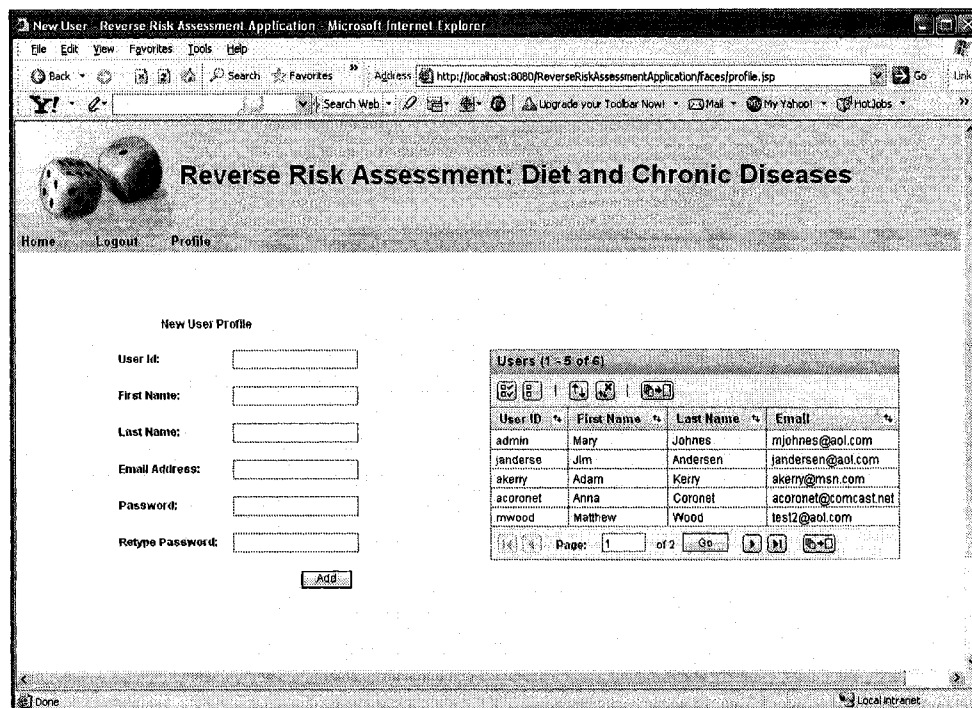


Figure 49: New User Page.

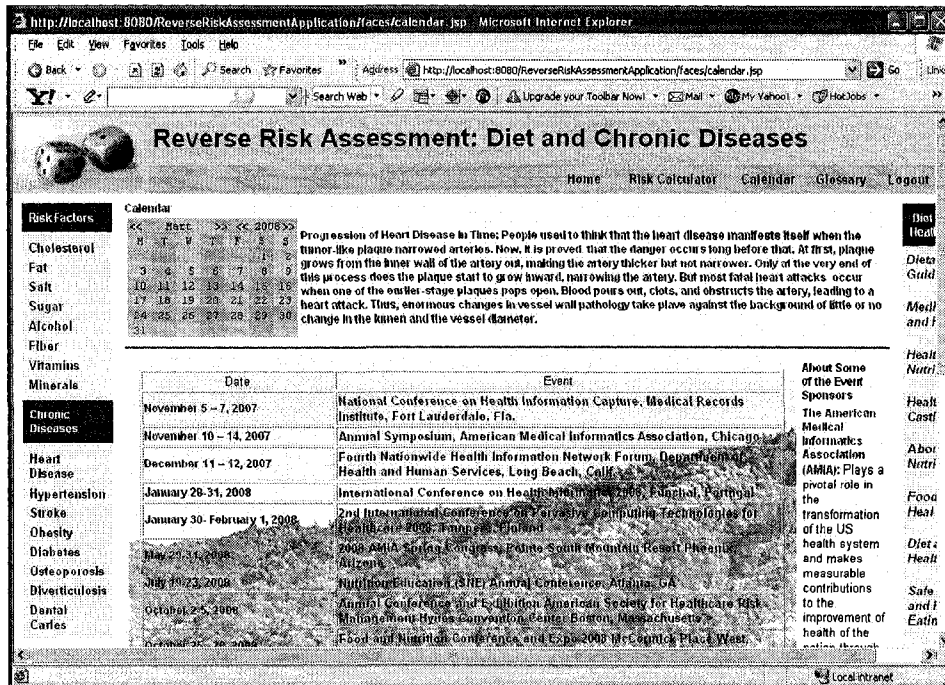


Figure 50: Calendar Page.

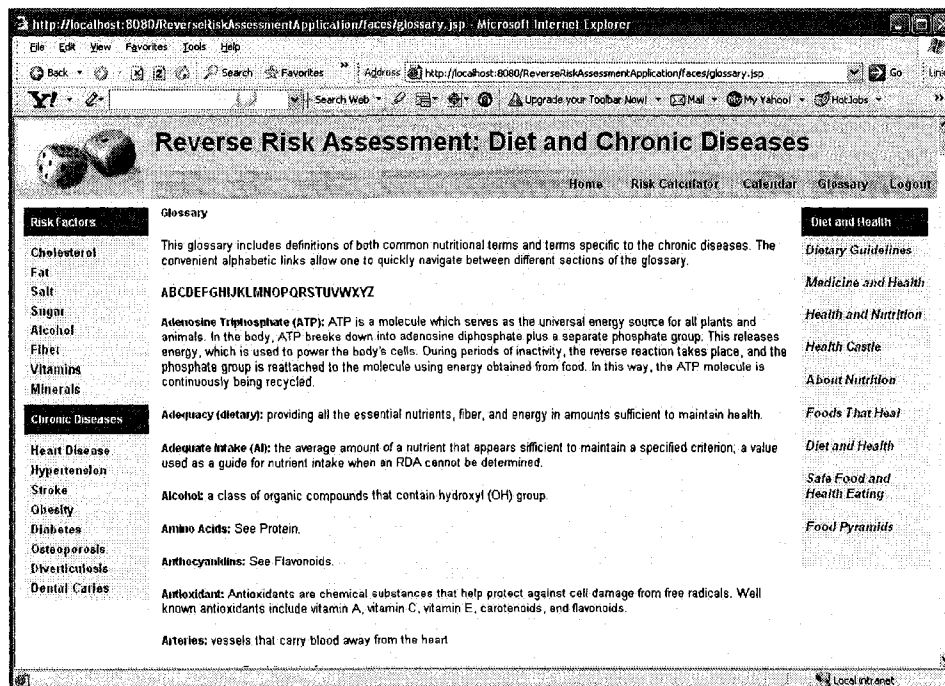


Figure 51: Glossary Page.

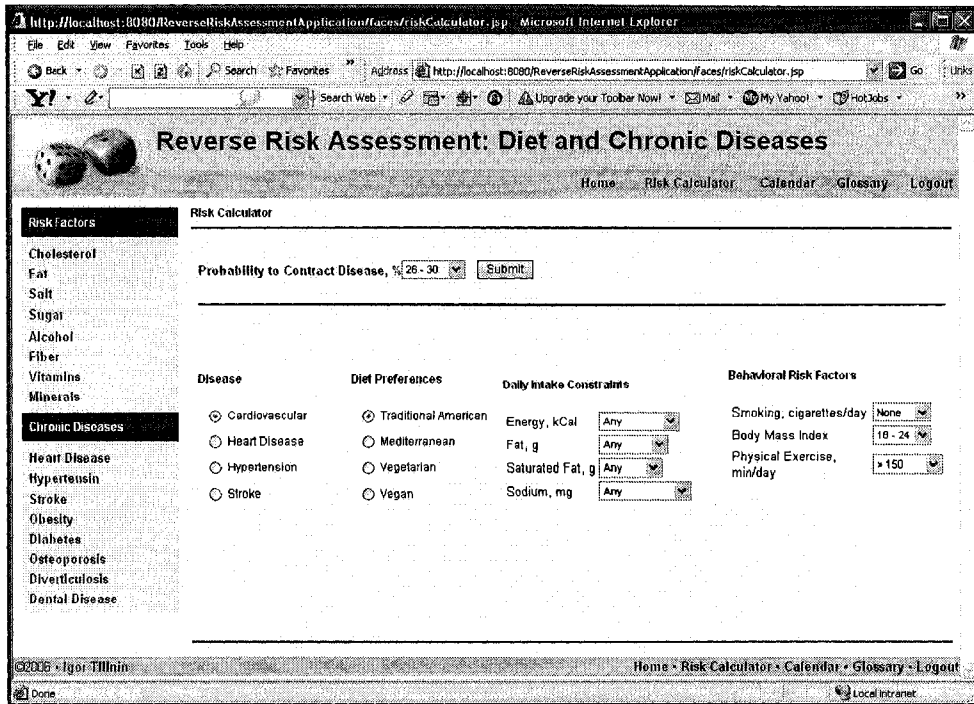


Figure 52: Risk Calculator Page.

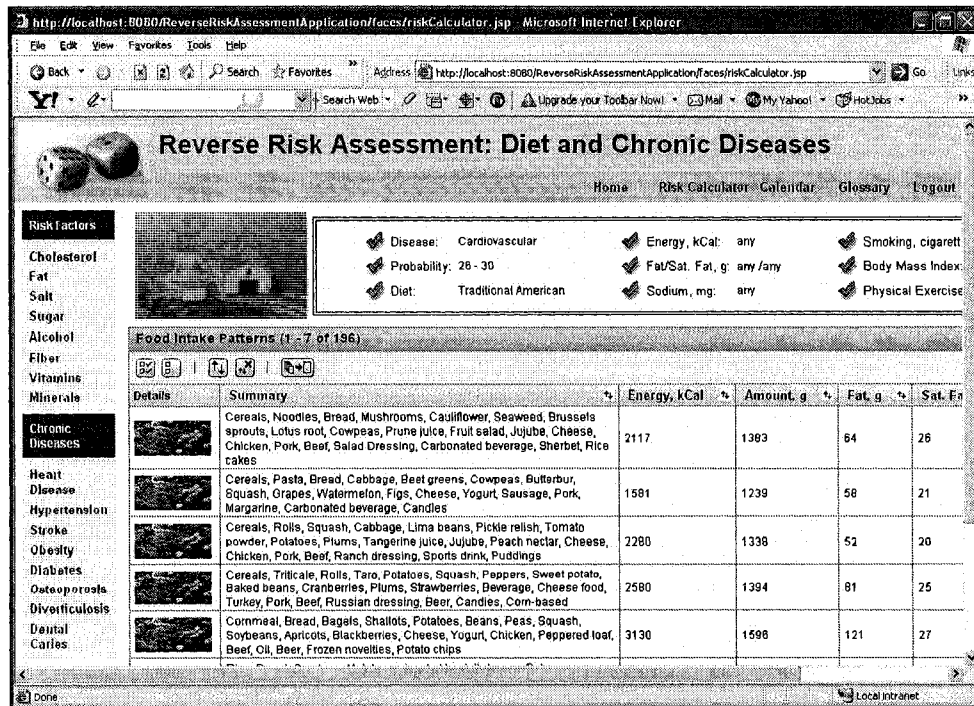


Figure 53: Search Results Page.

informational pages for the remaining diet-related risk factors and chronic diseases. Besides that, there 16 pages devoted to explanations of frequently asked questions. Thus, the total number of web pages in the application is 40.

3.5.2 Applet/JSF Integration

As already mentioned, the “Calendar” page includes a calendar applet. Incorporation of the applet pursued two objectives. First, it is convenient to have this utility when planning for events and browsing conference related information on the web. Secondly, combining an applet with JSF and making the applet work was an interesting task on its own. Being the official JSP standard JavaServer Faces are expected to work with applets seamlessly. Since there are many calendar applets available for free on the Internet, it was decided to use an out-of-the-shelf applet instead of building a new one from scratch. The applet used in this application was developed by Eric Bonder and downloaded from the web site [68]. The implementation of the “Calendar” page provides an example of how applets can be embedded into the JSF code. Table 7 shows the basic applet definition elements inside the HTML table cell tags in the “calendar.jsp” page. The applet definition starts with the element `<applet>` which several important attributes. The attribute “code” points to the applet class which is `calendar.class`. The “codebase” attribute value represents the directory name where the `calendar.class` file is located. In this particular case the directory name is “calendar” and it is located immediately beneath the root folder of the Reverse Risk Assessment application (See Fig. 3-2). In addition to attributes, the element `<applet>` has several child elements called `<param>`. Those elements contain names and

values of the applet parameters. For the sake of space and clarity, the actual names and values of the applet parameters are omitted in Table 7.

Table 7: Applet Definition and Configuration in JSP.

```
<td style="height: 128" width="182">
  <applet code="calendar.class" codebase="calendar" height="128" width="182">
    <param name="..." value="..." />
    .....
    <param name="..." value="..." />
  </applet>
</td>
```

3.5.3 Food Intake Pattern Display

The page “Search Results” (See Fig 53) contains a table with a summary of food items characterizing each pattern and several important parameters like energy, fat and sodium intake. However, the full pattern description need to include not only a list of food items but also details on the nutrient composition of each food item in the pattern. To display this kind of information in a convenient form, the results table provides an image link to the pattern XML file for each food intake pattern. Upon clicking the image to the left of the food intake summary, the browser loads the pattern XML file and opens it in the HTML format in a separate window. The document type definitions for the pattern XML

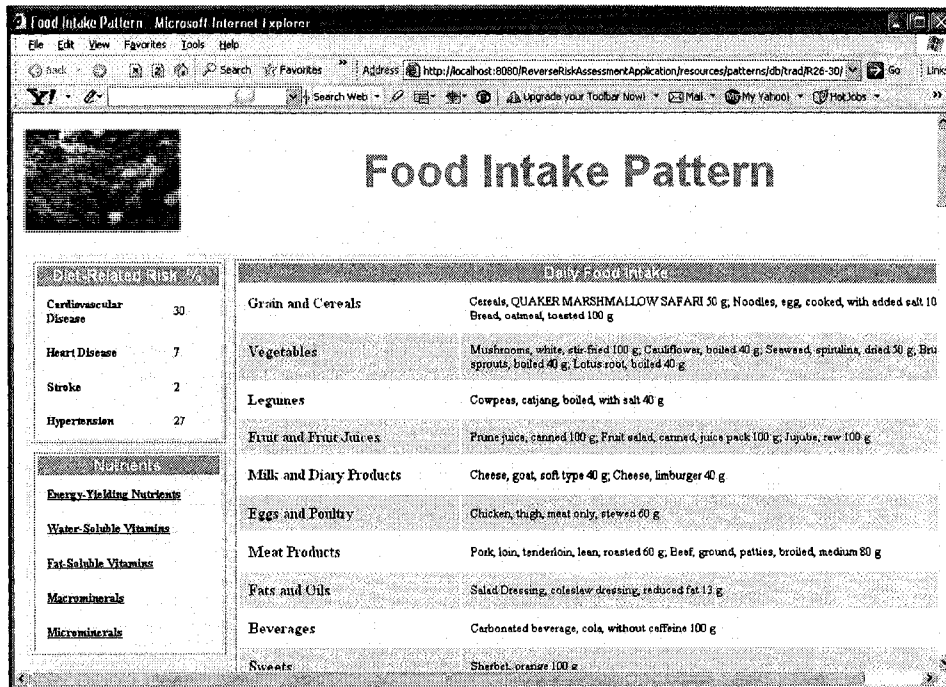


Figure 54: Top Part of Food Intake Pattern Page.

code are provided by the DTD file “FoodIntakePattern.dtd” (See Appendix B). Transformation of the XML into HTML format is performed at run time according to the style definitions and XSL transformations defined in “FoodIntakePattern.css” and “FoodIntakePattern.xsl” stylesheets, respectively. Appendix C contains the FoodIntakePattern.xsl stylesheet code. Fig. 54 displays the top part of the Food Intake Pattern page. One can see the main table with the food intake details and two smaller tables on the left hand side. One of those tables contains information about the probabilities to contract a chronic disease. It should be emphasized that the probability values displayed on the “Food Intake Pattern” page account only for diet risk factors associated with the pattern considered.

Food Intake Pattern - Microsoft Internet Explorer

Address: http://localhost:8080/ReverseRiskAssessmentApplication/resources/patterns/db/trad/R96-100/

Energy-Yielding Nutrients

Food Item	Quantity, g	Energy, kCal	Protein, g	Fat, g	Sat Fat, g	Trans Fat, g	Chol, mg	Carbo, g	Fibe, g
Macaroni, vegetable, cooked, enriched	100	128	4.53	0.11	0.02	0	0	26.61	4.3
Crackers, standard snack-type, sandwich, with cheese filling	40	190.8	3.72	8.44	2.43	0	0.8	24.68	0.76
Bread, rye	100	238	8.5	3.3	0.63	0	0	48.3	5.8
Horseradish-tees, pods, raw	100	37	2.1	0.2	0.03	0	0	8.53	3.2
Peas, green, boiled	40	33.6	2.14	0.09	0.02	0	0	6.26	2.2
Tomatoes, crushed, canned	100	32	1.64	0.28	0.04	0	0	7.29	1.9
Mushrooms, boiled	40	11.2	0.87	0.19	0.02	0	0	2.12	0.88
Sauerkraut, canned	100	19	0.91	0.14	0.03	0	0	4.28	2.9
Beans, navy, boiled, without salt	40	56	3.29	0.25	0.04	0	0	10.42	4.2
Papayas, raw	100	39	0.61	0.14	0.04	0	0	9.81	1.8
Peaches, dehydrated (low-moisture), sulfured, uncooked	100	325	4.89	1.03	0.11	0	0	83.18	0
Cheese, mexican, queso asadero	40	142.4	9.04	11.3	7.18	0	42	1.15	0
Cheese, mozzarella, whole milk	40	127.2	8.64	9.86	6.22	0	35.6	0.99	0
Chicken, cornish game hens, meat only, roasted	60	80.4	13.98	2.32	0.59	0	63.6	0	0
Luncheon Loaf, spiced	100	234	13.5	16.9	5.37	0	67	7	0
Beef, brain, simmered	80	120.8	9.34	8.42	1.92	0.41	2480	1.18	0
Fish oil, salmon	40	360.8	0	0	7.95	0	194	0	0
Wine, red, Zinfandel	100	89	0.07	0	0	0	0	2.86	0
Baking chocolate, mexican, squares	30	127.8	1.09	4.68	2.38	0	0	23.22	1.2
Breakfast bars, oats, sugar, raisins, coconut	30	139.2	2.94	5.28	3.81	0	0	20.01	0.93
Total:	1380	2551.2	91.8	112.93	39.05	0.41	2883	287.89	30.1
Recommended Daily Allowance (RDA)			58	65	20	0	<300	300	25
Tolerable Range			25 - 139	20 - 99	0 - 30	0 - 6	0 - 490	225 - 400	10 - 45

Figure 55: Energy-Yielding Nutrients Table - Food Intake Pattern Page.

Food Intake Pattern - Microsoft Internet Explorer

Address: http://localhost:8080/ReverseRiskAssessmentApplication/resources/patterns/db/trad/R96-100/

Water-Soluble Vitamins

Food Item	Quantity, g	C, mg	B1, mg	B2, mg	B3, mg	B6, mg	B12, µg	Folate, µg
Macaroni, vegetable, cooked, enriched	100	0	0.11	0.06	1.07	0.02	0	40
Crackers, standard snack-type, sandwich, with cheese filling	40	0.04	0.18	0.27	1.51	0.02	0.04	40
Bread, rye	100	0.4	0.43	0.33	3.8	0.07	0	110
Horseradish-tees, pods, raw	100	141	0.03	0.07	0.62	0.12	0	44
Peas, green, boiled	40	5.68	0.1	0.06	0.81	0.09	0	25.2
Tomatoes, crushed, canned	100	9.2	0.07	0.05	1.22	0.15	0	13
Mushrooms, boiled	40	1.6	0.03	0.12	1.78	0.04	0	7.2
Sauerkraut, canned	100	14.7	0.02	0.02	0.14	0.13	0	24
Beans, navy, boiled, without salt	40	0	0.09	0.03	0.26	0.06	0	56
Papayas, raw	100	61.8	0.03	0.03	0.34	0.02	0	38
Peaches, dehydrated (low-moisture), sulfured, uncooked	100	10.6	0.34	0.11	4.82	0.16	0	7
Cheese, mexican, queso asadero	40	0	0.01	0.09	0.07	0.02	0.4	3.2
Cheese, mozzarella, whole milk	40	0	0.31	0.11	0.04	0.02	0.29	3.2
Chicken, cornish game hens, meat only, roasted	60	0.36	0.24	0.14	3.76	0.21	0.18	1.2
Luncheon Loaf, spiced	100	0	0	0	0	0	0	0
Beef, brain, simmered	80	8.4	0.06	0.17	2.9	0.11	8.08	4
Fish oil, salmon	40	0	0	0	0	0	0	0
Wine, red, Zinfandel	100	0	0	0	0	0	0	0
Baking chocolate, mexican, squares	30	0.03	0.02	0.03	0.55	0.01	0	1.5
Breakfast bars, oats, sugar, raisins, coconut	30	0.3	0.08	0.03	0.52	0.1	0	24.3
Total:	1380	254.11	1.37	1.72	24.21	1.35	8.99	466.8
Recommended Daily Allowance (RDA)		60	1.5	1.7	20	2	6	400
Tolerable Range		30 - 2000	0.8 - n/a	0.9 - n/a	10 - 35	1 - 100	3 - n/a	200 - 1000

Figure 56: Water-Soluble Vitamins Table - Food Intake Pattern Page.

Food Intake Pattern - Microsoft Internet Explorer

Address: http://localhost:8080/ReverseRiskAssessmentApplication/resources/patterns/db/trad/R96-100/

Macrominerals						
Food Item	Quantity, g	Ca, mg	Na, mg	K, mg	P, mg	Mg, mg
Macaroni, vegetable, cooked, enriched	100	11	6	31	50	19
Crackers, standard snack-type, sandwich, with cheese filling	40	102.8	560.4	171.6	162.4	14.4
Bread, rye	100	73	660	166	125	40
Horseradish-tee, pods, raw	100	30	42	461	30	45
Peas, green, boiled	40	10.8	1.2	108.4	46.8	15.6
Tomatoes, crushed, canned	100	34	132	293	32	20
Mushrooms, boiled	40	2.4	0.8	142.4	34.8	4.8
Sauerkraut, canned	100	30	661	170	20	13
Beans, navy, boiled, without salt	40	27.6	0	155.6	57.6	21.2
Papayas, raw	100	24	3	257	5	10
Peaches, dehydrated (low-moisture), sulfured, uncooked	100	38	10	1351	162	57
Cheese, mexican, queso asadero	40	264.4	262	34.4	177.2	10.4
Cheese, mozzarella, whole milk	40	230	166	30	164.8	8.4
Chicken, cornish game hens, meat only, roasted	60	7.8	37.8	130	89.4	11.4
Luncheon Loaf, spiced	100	109	1226	270	195	24
Beef, brain, simmered	80	7.2	86.4	195.2	268	9.6
Fish oil, salmon	40	0	0	0	0	0
Wine, red, Zinfandel	100	0	0	0	0	0
Baking chocolate, mexican, squares	30	10.2	0.9	119.1	42.6	28.5
Breakfast bars, oats, sugar, raisins, coconut	30	18	83.4	97.8	83.1	30.3
Total:	1380	1030.2	3938.9	4203.5	1745.7	382.6
Recommended Daily Allowance (RDA)		1000	2400	2000	1000	400
Tolerable Range		500 - 2500	500 - 10000	1000 - 10000	500 - 4000	200 - 800

Figure 57: Fat-Soluble Vitamins Table - Food Intake Pattern Page

Food Intake Pattern - Microsoft Internet Explorer

Address: http://localhost:8080/ReverseRiskAssessmentApplication/resources/patterns/db/trad/R96-100/

Microminerals						
Food Item	Quantity, g	Fe, mg	Zn, mg	Cu, mg	Mn, mg	Se, µg
Macaroni, vegetable, cooked, enriched	100	0.49	0.44	0.09	0.99	19.8
Crackers, standard snack-type, sandwich, with cheese filling	40	0.96	0.25	0.03	0.12	8.6
Bread, rye	100	2.83	1.14	0.19	0.82	30.9
Horseradish-tee, pods, raw	100	0.26	0.45	0.08	0.26	0.7
Peas, green, boiled	40	0.62	0.48	0.07	0.21	0.76
Tomatoes, crushed, canned	100	1.3	0.27	0.18	0.18	0.6
Mushrooms, boiled	40	0.7	0.35	0.2	0.05	4.76
Sauerkraut, canned	100	1.47	0.19	0.1	0.15	0.6
Beans, navy, boiled, without salt	40	0.94	0.41	0.08	0.21	1.16
Papayas, raw	100	0.1	0.07	0.02	0.01	0.6
Peaches, dehydrated (low-moisture), sulfured, uncooked	100	5.51	0.78	0.49	0.41	0
Cheese, mexican, queso asadero	40	0.2	1.21	0.01	0.01	5.8
Cheese, mozzarella, whole milk	40	0.08	0.98	0.01	0	6.44
Chicken, cornish game hens, meat only, roasted	60	0.46	0.92	0.04	0.01	12.48
Luncheon Loaf, spiced	100	1.35	1.95	2.31	0	0
Beef, brain, simmered	80	1.84	0.87	0.18	0.02	17.44
Fish oil, salmon	40	0	0	0	0	0
Wine, red, Zinfandel	100	0	0	0	0	0
Baking chocolate, mexican, squares	30	0.65	0.38	0.2	0.13	0.54
Breakfast bars, oats, sugar, raisins, coconut	30	0.95	0.48	0.11	0	4.8
Total:	1380	28.81	11.62	4.39	3.58	115.98
Recommended Daily Allowance (RDA)		18	15	2	2	70
Tolerable Range		18 - 40	18 - 2000	1 - 4	1 - 10	35 - 1000

Figure 58: Microminerals Table - Food Intake Pattern Page.

Below the “Diet-Related Risk” table, there is a table with quick navigation links to the 5 tables with the information on nutritional composition of the pattern food items. These tables are shown in Fig. 55-58 and include energy-yielding nutrients, water-soluble vitamins, fat-soluble vitamins, macrominerals, and microminerals. For reference purposes, the tables display the total amount of a certain nutrient, its Recommended Daily Allowance (RDA) and tolerable range for every pattern.

This concludes the overview of the GUI implementation. The complete listing of the application JSP pages is presented in Table 8.

3.5.4 Images

A vast majority of images used in the application were created by the author. A logo image was made partially using a free thumbnail image available on the Internet. All images were embedded into the JSP pages exclusively for composition and layout purposes only.

3.6 Java Beans

The application business logic is encapsulated in Java classes which are packaged into one package “reverseriskassessment”. The “reverseriskassessment” package consists of 43 classes including 40, the so-called, “backing beans” and 3 default managed beans. The “backing beans” are Java Beans pertaining to each JSP page in the application. The backing beans are generated based on the respective JSP layout and support all widget

functionality on a page. The remaining three classes are application bean, session bean, and request bean. The default managed beans play a special role in the application. They are instantiated by the application and are stored in the application, session and request scopes, respectively.

Table 8: Listing of Application Package Java Classes and JSP

No	Filename	JSP Supported	Description
1	alcohol.java	alcohol.jsp	Automatically generated based on JSP layout.
2	ApplicationBean1.java	n/a	Managed bean with application scope.
3	calendar.java	calendar.jsp	Automatically generated based on JSP layout.
4	cholesterol.java	cholesterol.jsp	Automatically generated based on JSP layout.
5	dentalDisease.java	dentalDisease.jsp	Automatically generated based on JSP layout.
6	diabetes.java	diabetes.jsp	Automatically generated based on JSP layout.
7	diverticulosis.java	diverticulosis.jsp	Automatically generated based on JSP layout.
8	faqAlcohol.java	faqAlcohol.jsp	Automatically generated based on JSP layout.
9	faqCholesterol.java	faqCholesterol.jsp	Automatically generated based on JSP layout.
10	faqDentalDisease.java	faqDentalDisease.jsp	Automatically generated based on JSP layout.

11	faqDiabetes.java	faqDiabetes.jsp	Automatically generated based on JSP layout.
12	faqDiverticulosis.java	faqDiverticulosis.jsp	Automatically generated based on JSP layout.
13	faqFats.java	faqFats.jsp	Automatically generated based on JSP layout.
14	faqFiber.java	faqFiber.jsp	Automatically generated based on JSP layout.
15	faqHeartDisease.java	faqHeartDisease.jsp	Automatically generated based on JSP layout.
16	faqHypertension.java	faqHypertension.jsp	Automatically generated based on JSP layout.
17	faqMinerals.java	faqMinerals.jsp	Automatically generated based on JSP layout.
18	faqObesity.java	faqObesity.jsp	Automatically generated based on JSP layout.
19	faqOsteoporosis.java	faqOsteoporosis.jsp	Automatically generated based on JSP layout.
20	faqSalt.java	faqSalt.jsp	Automatically generated based on JSP layout.
21	faqStroke.java	faqStroke.jsp	Automatically generated based on JSP layout.
22	faqSugar.java	faqSugar.jsp	Automatically generated based on JSP layout.
23	faqVitamins.java	faqVitamins.jsp	Automatically generated based on JSP layout.
24	fat.java	fat.jsp	Automatically generated based on JSP layout.
25	fiber.java	fiber.jsp	Automatically generated based on

			JSP layout.
26	glossary.java	glossary.jsp	Automatically generated based on JSP layout.
27	heartDisease.java	heartDisease.jsp	Automatically generated based on JSP layout.
28	hypertension.java	hypertension.jsp	Automatically generated based on JSP layout.
29	index.java	index.jsp	Automatically generated based on JSP layout.
30	login.java	login.jsp	Contains authentication logic and opens a session.
31	minerals.java	minerals.jsp	Automatically generated based on JSP layout.
32	newUser.java	newUser.jsp	Creates a new user profile in the database.
33	obesity.java	obesity.jsp	Automatically generated based on JSP layout.
34	osteoporosis.java	steoporosis.jsp	Automatically generated based on JSP layout.
35	profile.java	profile.jsp	Updates a user profile in the database.
36	RequestBean1.java	n/a	Managed bean with request scope.
37	riskCalculator.java	riskCalculator.jsp	Makes initial processing of the search parameters and passes the request to the searchResults bean.
38	salt.java	salt.jsp	Automatically generated based on JSP layout.
39	searchResults.java	searchResults.jsp	Builds the actual query and populates the results table.

40	sessionBean1.java	n/a	Managed bean with session scope.
41	stroke.java	stroke.jsp	Automatically generated based on JSP layout.
42	sugar.java	sugar.jsp	Automatically generated based on JSP layout.
43	vitamins.java	vitamins.jsp	Automatically generated based on JSP layout.

The “Description” column in Table 8 provides brief explanation of the bean functionality. Those beans which have been generated by the IDE as a result of JSP changes only and have not been customized separately are marked as “Automatically generated based on JSP layout.”

Most important customizations are made to the riskCalculator.java and searchResults.java beans. In particular, the risk criteria entered on the riskCalculator.jsp page are analyzed and the total probability of contracting the selected disease is used to determine the contribution of diet-related risk factors. The diet-related, or partial, probability is then passed over along with other parameters to the searchResults.java bean which performs the actual querying of the pattern database.

3.7 System Specifications

This section summarizes the implemented system specifications in terms of hardware and software components, which are conveniently grouped into Tables 9 and 10, respectively. The system uses Apache Tomcat version 5.0.28 and MySQL relational database

management system (RDBMS) version 4.0.11 and can be deployed on a Windows workstation or server.

Table 9: System Hardware Specifications.

No	Hardware Component	Description
1	Processor	Intel T2250, 1.73 GHz
2	Hard Drive	80 GB
3	Memory (RAM)	2 GB

Table 10: System Software Specifications.

No	Software Component	Description
1	Operating System	Windows XP SP2
2	Apache Tomcat Application Server	5.0.28
3	My SQL RDBMS	4.1.11 (Windows)

3.8 Application Deployment

The application deployment procedure encompasses the following major steps: (1) Generation of the web archive file and copying the web archive to the Apache Tomcat webapps directory, (2) Modifying the application web.xml file and configuring a listener

class, (3) Downloading and installing standard libraries on Apache Tomcat, (4) Configuring database connection and data source. All these steps must be performed carefully to ensure proper application deployment. As simple as it seems, the deployment procedure sometimes takes more time than expected due to potential pitfalls. To ensure repeatability and application portability, the procedure has been documented is described in detail below.

3.8.1 Generation of Web Archive File

The Java Studio Creator IDE allows the developer automatically generate a web archive.

The procedure of generating a web archive in Creator is as follows:

1. Open the project in the Creator IDE:
2. In the Project pane, select ReverseRiskAssessmentApplication and right-click to display a submenu of options. Seelect “Export WAR File” from the drop down menu.
3. Export WAR File dialog opens.
4. At the prompt select JSEE 1.4 and specify the destination file:
“C:\Program Files\Apache Software Foundation\Tomcat
5.0\webapps\ReverseRiskAssessmentApplication”
5. Click “OK”. The IDE generates a web archive and copies it into the Apache Tomcat webapps directory.

3.8.2 Application Listener Class Configuration

When running the Apache Tomcat 5.0.28 on Windows, it is necessary to configure the listener class. This is done by following the steps:

1. Restart Tomcat to expand the web archive into the directory structure. Open the web.xml file located in the WEB-INF directory beneath the application root directory.
2. Edit the web.xml file by adding the code as shown in Table 11.
3. Save the web.xml file and restart Apache Tomcat for the changes to take effect.

The next step is to install some libraries which are necessary to run a JSF application.

Table 11: Listener Class Configuration Code.

```
<web-app>
.....
<listener>
  <listener-class>com.sun.faces.config.ConfiguredListener</listener-class>
</listener>
</web-app>
```

3.8.3 Standard Libraries Installation

The standard libraries for running a JSF application on Apache Tomcat can be downloaded from the Apache web site at the URL specified below:

http://jakarta.apache.org/site/downloads/downloads_taglibs-standard.cgi

The page contains a link to the jackarta-taglibs-standard-1.1.2.zip archive, which should be downloaded and unzipped in a temporary directory. After expanding the archive a directory structure is created with a lib folder beneath the root directory. The lib folder contains two jar files jstl.jar and standard.jar which have to be copied to the Tomcat “common\lib” folder. Once the libraries are installed the Tomcat needs to be restarted.

Finally, it is necessary to configure the database connection. This procedure is explained in the next section.

3.8.4 Database Connection Configuration

It is assumed that an instance of the MySQL is installed and running on the Apache Tomcat host machine and the MySQL-JDBC connector archive file (mysql-connector-java-3.2.0-alpha-bin.jar) is available. The examples below also assume that there is a database “pca” which contains all the necessary data tables. The database connection configuration procedure consists of the following steps:

1. Copy the connector jar file “mysql-coonector-java-3.2.0-alpha-bin.jar” into the Tomcat “common\lib” directory.
2. Navigate to the Tomcat “config” folder and open the “server.xml” file.

3. Locate the element `<Context>` in the “server.xml” and add the following code as shown in Table 12:

Table 12: Database Connection Configuration.

```
<!-- Reverse Risk Assessment Database connection -->
<Context path="/ReverseRiskAssessmentApplication"
docBase="ReverseRiskAssessmentApplication" debug="5" reloadable="true"
crossContext="true">
  <Resource name="jdbc/RiskAssessment" auth="Container"
    type="javax.sql.DataSource"
    description="Risk Assessment Database">
  </Resource>
  <ResourceParams name="jdbc/RiskAssessment">
    <parameter>
      <name>factory</name>
      <value>org.apache.commons.dbcp.BasicDataSourceFactory</value>
    </parameter>
    <parameter>
      <name>validationQuery</name>
      <value>select * from pattern;</value>
    </parameter>
    <parameter>
```

```
<name>maxWait</name>

<value>5000</value>

</parameter>

<parameter>

<name>maxActive</name>

<value>4</value>

</parameter>

<parameter>

<name>password</name>

<value>pca</value>

</parameter>

<parameter>

<name>url</name>

<value>jdbc:mysql://localhost/risk_assessment?autoReconnect=true</value>

</parameter>

<parameter>

<name>driverClassName</name>

<value>org.gjt.mm.mysql.Driver</value>

</parameter>

<parameter>

<name>maxIdle</name>

<value>2</value>
```

```
</parameter>
<parameter>
  <name>username</name>
  <value>pca</value>
</parameter>
</ResourceParams>
</Context>
```

4. Save the server.xml file and restart the Tomcat.

Note that there is no port number in the connection string specified in the parameter “url” in contrast to some official documentation recommendations.

This concludes the deployment procedure for the Reverse Risk Assessment application.

3.9 Food Intake Pattern Generation

The Reverse Risk Assessment application relies on the database of food intake patterns which are generated in a certain format convenient for data display and retrieval. However, irrespective of the presentation layer, the patterns need to be generated taking into account real food consumption trends to make risk assessment meaningful. The current application implementation differentiates food intake patterns according to 4 major categories or diets: (a) American Traditional, (b) Mediterranean, (c) Vegetarian,

and (d) Vegan. It is also implied that each pattern is characterized by a variety of food items from 17 major food groups. The food items, in turn, are described by their nutritional value in terms of energy-yielding nutrients, vitamins and minerals. Thus, the process of pattern generation is supposed to produce pattern data objects which are characterized by a large number of parameters.

3.9.1 Raw Data Preparation

Although, the food composition databases are available [56], they are not directly suited for food intake pattern generation due to several reasons. First of all, the food names in the existing databases are sometimes too long, carrying details about food preparation. Secondly, the processed food items often have names which have little meaning to a consumer and serve as company's advertising labels rather than provide reasonable description of food. Thirdly, it has been found that some database don't contain important information on the trans fat content, and finally, thorough examination of a large number of food items in these databases uncovered a significant number of typos. In this project, the Food and Drug Administration database entries [56] are taken as a base line for creating a raw version of food composition data. The long food names were abbreviated without losing their descriptive nature and typos were corrected. The food item information was written into two text files which could be quickly and efficiently processed by a Perl script. One of them contains food IDs and food names, whereas the other stores data on nutrient content. The two datasets are related by the food ID key.

3.9.2 Pattern Generation Script

Analysis of the food intake data structure shows that a program written in Perl would provide the quickest solution for pattern generation. The Perl program called “make_fip.pl” has about 1500 lines of code and takes 2 minutes on average to generate a food intake pattern for a 2.0 GHz processor. The generation process is based on random selection of food items from different food groups in accordance to prevalence of those items in a given diet. The data selected from the raw database (See the previous section) are then used to populate the data structure and fill in the pattern template described below.

3.9.3 Pattern Data Structure

The pattern data structure is a key element in the pattern generation program. The data structure plays a role of the pattern class and accounts for all necessary food intake pattern characteristics including total food and calorie intake, as well as food composition. The data structure code has been written in Perl and is shown in Table 13. As it follows from Table 13, the food intake data structure has 3 major properties: (a) foodGroups, (b) nutrients, and (c) properties. Each of those properties has its own substructure to hold data necessary for the pattern definition.

Table 13: Food Intake Pattern Data Structure Code.

```
my %foodIntakePattern = (  
    foodGroups => [ {name=>"TBD", composition=>"Empty",
```



```

foodItemsTotal=>0,

        foodItems => [ { itemID=>"00000", name=>"TBD",
group=>"FoodGroup", amount=>100,

        itemNutrients=>{ "Energy"=>0, "Protein"=>0, "Fat"=>0,
"Sat_Fat"=>0, "Trans_Fat"=>0, "Chol"=>0, "Carbs"=>0, "Fiber"=>0, "Energy"=>0,
                "C"=>0, "B1"=>0, "B2"=>0, "B3"=>0,
"B6"=>0, "B12"=>0, "Folate"=>0,
                "A"=>0, "D"=>0, "E"=>0, "K"=>0,
                "Ca"=>0, "Na"=>0, "min_K"=>0, "P"=>0,
"Mg"=>0,
                "Fe"=>0, "Zn"=>0, "Cu"=>0, "Mn"=>0, "Se"=>0
                }
        }
    ],

    nutrients => [
        { ID=>"208", name=>"Energy",   group=>"Energy-Yielding
Nutrients", unit=>"kCal", amount=>0 },
        { ID=>"203", name=>"Protein",   group=>"Energy-Yielding
Nutrients", unit=>"g", amount=>0 },
        { ID=>"204", name=>"Fat",       group=>"Energy-Yielding

```

```

Nutrients", unit=>"g", amount=>0 },
        { ID=>"606", name=>"Sat Fat",  group=>"Energy-Yielding
Nutrients", unit=>"g", amount=>0 },
        { ID=>"605", name=>"Trans Fat",  group=>"Energy-
Yielding Nutrients", unit=>"g", amount=>0 },
        { ID=>"601", name=>"Chol",  group=>"Energy-Yielding
Nutrients", unit=>"mg", amount=>0 },
        { ID=>"205", name=>"Carbs",  group=>"Energy-Yielding
Nutrients", unit=>"g", amount=>0 },
        { ID=>"291", name=>"Fiber",  group=>"Energy-Yielding
Nutrients", unit=>"g", amount=>0 },

        { ID=>"401", name=>"C",  group=>"Water-Soluble
Vitamins", unit=>"mg", amount=>0 },
        { ID=>"404", name=>"B1",  group=>"Water-Soluble
Vitamins", unit=>"mg", amount=>0 },
        { ID=>"405", name=>"B2",  group=>"Water-Soluble
Vitamins", unit=>"mg", amount=>0 },
        { ID=>"406", name=>"B3",  group=>"Water-Soluble
Vitamins", unit=>"mg", amount=>0 },
        { ID=>"415", name=>"B6",  group=>"Water-Soluble
Vitamins", unit=>"mg", amount=>0 },

```

```
        { ID=>"418", name=>"B12",  group=>"Water-Soluble
Vitamins", unit=>"mcg", amount=>0 },
        { ID=>"417", name=>"Folate", group=>"Water-Soluble
Vitamins", unit=>"mcg", amount=>0 },
        { ID=>"320", name=>"A",  group=>"Fat-Soluble Vitamins",
unit=>"mcg", amount=>0 },
        { ID=>"324", name=>"D",  group=>"Fat-Soluble Vitamins",
unit=>"mcg", amount=>0 },
        { ID=>"323", name=>"E",  group=>"Fat-Soluble Vitamins",
unit=>"mg", amount=>0 },
        { ID=>"430", name=>"K",  group=>"Fat-Soluble Vitamins",
unit=>"mcg", amount=>0 },

        { ID=>"301", name=>"Ca", group=>"Macrominerals",
unit=>"mg", amount=>0 },
        { ID=>"307", name=>"Na", group=>"Macrominerals",
unit=>"mg", amount=>0 },
        { ID=>"306", name=>"K",  group=>"Macrominerals",
unit=>"mg", amount=>0 },
        { ID=>"305", name=>"P",  group=>"Macrominerals",
unit=>"mg", amount=>0 },
        { ID=>"304", name=>"Mg", group=>"Macrominerals",
```

```

unit=>"mg", amount=>0 },
      { ID=>"303", name=>"Fe", group=>"Microminerals",
unit=>"mg", amount=>0 },
      { ID=>"309", name=>"Zn", group=>"Microminerals",
unit=>"mg", amount=>0 },
      { ID=>"312", name=>"Cu", group=>"Microminerals",
unit=>"mg", amount=>0 },
      { ID=>"315", name=>"Mn", group=>"Microminerals",
unit=>"mg", amount=>0 },
      { ID=>"317", name=>"Se", group=>"Microminerals",
unit=>"mcg", amount=>0 },

    ],

    foodGroups => [
      { ID=>"01", name=>"Grain and Cereals",
composition=>"Empty", members=>[0] },
      { ID=>"02", name=>"Vegetables",
composition=>"Empty", members=>[0] },
      { ID=>"03", name=>"Legumes",
composition=>"Empty", members=>[0] },
      { ID=>"04", name=>"Fruit and Fruit Juices",

```

```

composition=>"Empty", members=>[0] },
        { ID=>"05", name=>"Milk and Dairy Products",
composition=>"Empty", members=>[0] },
        { ID=>"06", name=>"Eggs and Poultry",
composition=>"Empty", members=>[0] },
        { ID=>"07", name=>"Meat Products",
composition=>"Empty", members=>[0] },
        { ID=>"08", name=>"Fish and Seafood",
composition=>"Empty", members=>[0] },
        { ID=>"09", name=>"Fats and Oils", composition=>"Empty",
members=>[0] },
        { ID=>"10", name=>"Beverages", composition=>"Empty",
members=>[0] },
        { ID=>"11", name=>"Sweets", composition=>"Empty",
members=>[0] },
        { ID=>"12", name=>"Nuts and Seeds",
composition=>"Empty", members=>[0] },
        { ID=>"13", name=>"Snacks", composition=>"Empty",
members=>[0] },
    ],

    properties => [ ID=>"00000", energy=>0, summary=>"", amount=>0,

```

```
dietType=>"Diet",  
  
        diseaseRisk=>{ "CVD"=>0, "HeartDisease"=>0,  
"Stroke"=>0, "Hypertension"=>0 },  
  
        energyRange=>"0 - 0", sodiumRange=>"0 - 0"  
  
        ]  
  
);
```

3.9.4 Food Intake Pattern Template

When generating a food intake XML file, the Perl program employs an XML template which contains an XML data structure to be populated with all the information about the food intake pattern. The template file is eleven-page long, therefore, for lack of space Fig. 59 shows only the XML element hierarchy. The template has the “FoodIntakePattern” element at the root level and a number of lower level children. The element multiplicity is denoted with the plus sign. Some of the elements have attributes as defined by the DTD in Appendix B.

During the food intake pattern generation process, several elements are populated with values obtained by summation of data stored in other elements. Thus, the elements “Quantity” and “TotalEnergy” are populated with the total food and energy intakes, respectively. Correspondingly, the element “AmountAvailable”, which is a child of the element “Nutrient”, is filled in with total amount of this particular nutrient in the whole pattern.

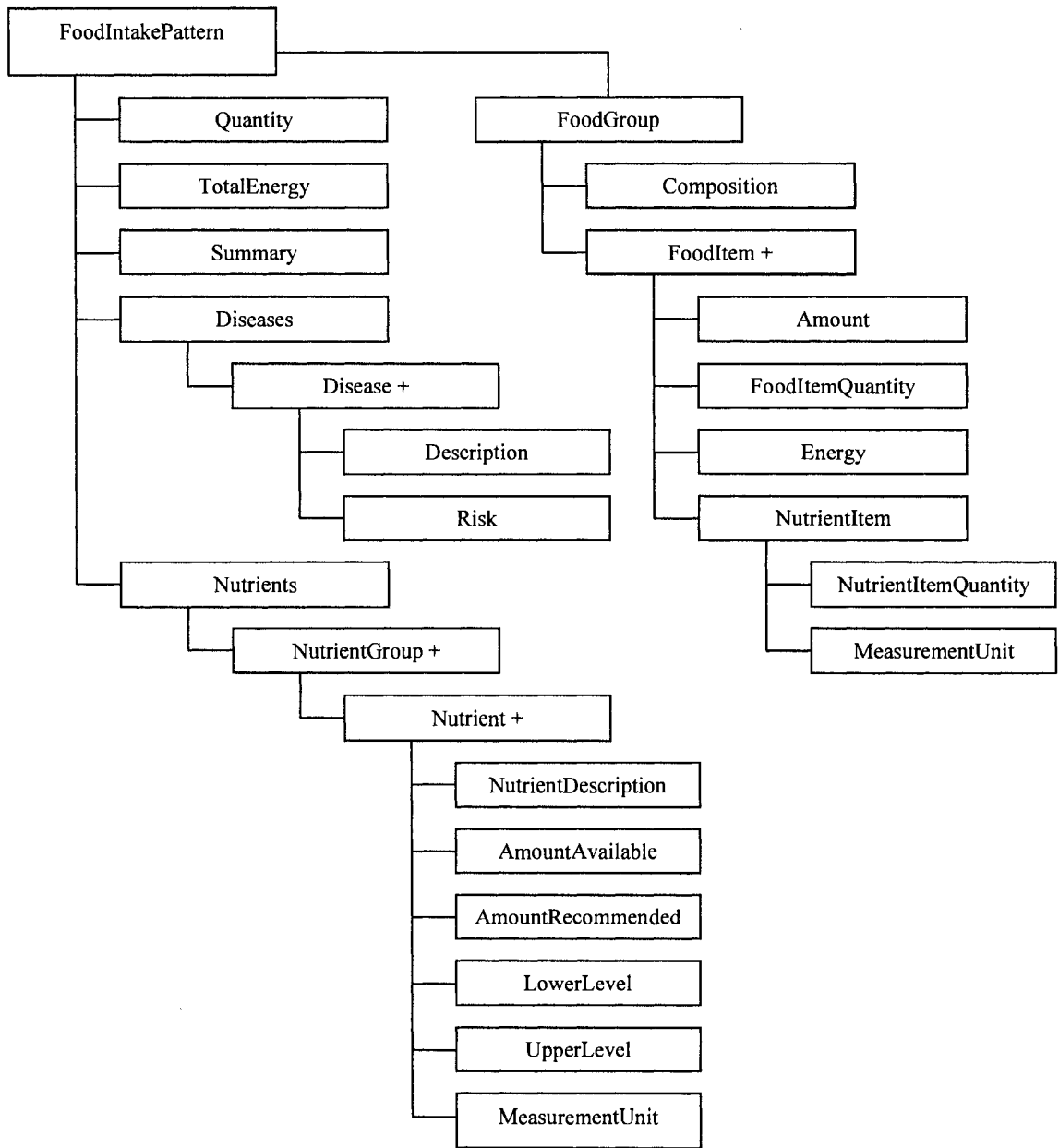


Figure 59: Schematic Representation of Food Intake Pattern XML Structure.

3.10 Diet Modeling

The Reverse Risk Assessment application offers four options for specifying a diet: (1) Traditional American, (2) Mediterranean, (3) Vegetarian, and (4) Vegan. Generation of food intake patterns pertaining to these diets should account for a number of constraints. First of all, each diet is characterized by a certain combination of foods. Secondly, diet variations are affected by traditions and availability of food items. This section provides diet definitions and discusses the algorithm for modeling food intakes corresponding to the aforesaid diet options.

3.10.1 Traditional American Diet

Diet is commonly defined as a regularly-consumed selection of food. Evidently, the term “regularly-consumed” implies an extended period of time over which the food set is used to satisfy nutritional needs. From the risk assessment point of view, this period should be comparable with the life span or, at least, with the generation which is estimated to span about 25–30 years. Below it is assumed that there is a set of basic food items which characterizes the traditional way Americans consume food on a regular basis and which doesn’t vary too much over the extended period of time of about 30-40 years.

Analysis of the U.S. food consumption data collected from 1909 to 2004 [69, 70] confirms this assumption. Indeed, although the notion of the so-called Traditional American diet should be understood dynamically, the basic set of major food items and their amounts consumed in America varied slowly over the period of time which spans

almost one century. For instance, the daily energy intake from meat, poultry and fish has changed insignificantly from 579 to 523 kCal since 1970. Consumption of another major staple component of the American diet, white potatoes, followed the same pattern. Over the period from 1970 to 2004, the energy intake for potatoes increased slightly from about 90 to 93 kCal on a daily basis. The same is true about consumption of vegetables and fruits, which showed slight increase of 2 and 25%, respectively. Though consumption of grain products dropped to its lowest point in 1969-1970, by the end of 2004 it approached the level observed in 1909-1919. In other words, it is fair to define the Traditional American diet as a staple food item set whose components can be derived from the food consumption data across the country over the last 30-40 years. The diet definition should include major food groups and their subgroups, as well as the amount of food consumed on average in each of those groups. The corresponding Traditional American diet definition is shown in Table 14. It is seen from Table 14 that the Traditional American diet is characterized by a significant intake of red meats, fats and dairy products.

3.10.2 Mediterranean Diet

Mediterranean diet started attracted attention from the media since early 1990-ties as researchers tried to explain a number of nutritional paradoxes related to the risk of contracting cardiovascular diseases as observed in the population of the countries around the basin of the Mediterranean Sea. In particular, it was found that the rate of mortality and morbidity in the countries like Greece, Italy, Israel, Lebanon, Spain, Tunisia, Turkey and some others was much lower than in the western countries despite the fact that the

total amount of fat consumed in the Mediterranean diet might account for 40% of all calorie intake. It should be stressed that the term Mediterranean diet is gradually emerged and now used as a collective notion for the best nutrition choices known in the countries of the Mediterranean region. In other words, it denotes a certain type of diet which combines together the healthiest eating habits of the people living in this part of the world.

When compiling data on the Mediterranean diet definition, the following characteristics of the Mediterranean diet have been taken into account [13, 71]:

- Relatively high consumption of healthy fats like olive and canola oils.
- Significant amount of bread and cereals in the diet.
- Abundant quantities of vegetables, fruits, and legumes.
- Regular consumption of cheese and yogurt.
- Moderate consumption of fish, poultry, and eggs.
- Drinking wine in moderation on a regular basis.
- Relatively high consumption of nuts and seeds.
- Very low consumption of red meats (few times a month).

Table 14 contains some quantitative details on food intakes for the Mediterranean diet. It follows from Table 14, that the Mediterranean diet is characterized by a big reduction in red meat, butter and margarine, on one hand, and significant increase in vegetables and fruits components, on the other hand, as compared to the Traditional American diet.

Table 14: Diet Type Definitions.

Food Group	Subgroup	Diet			
		Trad. American	Mediterranean	Vegetarian	Vegan
Meat and Meat Products	Beef	80	5	0	0
	Pork	60	5	0	0
	Lamb	2	10	0	0
	Sausages	40	0	0	0
Diary Products	Cheese	40	25	30	0
	Milk/Yogurt	200	200	200	0
Eggs and Poultry	Chicken/Turkey	60	20	0	0
	Eggs	40	15	40	0
Fish/Seafood	n/a	15	40	0	0
Grains and Cereals	Cereals	50	120	150	150
	Bread	100	200	150	150
	Breakfast	100	100	100	100
Vegetables	n/a	500	750	750	850
Legumes	n/a	20	40	50	100
Fruits and Juices	n/a	370	400	400	450
Nuts and Seeds	n/a	5	20	30	40
Fats and Oils	Butter	10	2	0	0
	Oil	30	40	30	20
	Margarine	20	0	0	0
Sweets	n/a	100	20	60	30
Beverages	n/a	200	200	200	200
Snacks	n/a	30	0	60	50

3.10.3 Vegetarian Diet

The Vegetarian [72, 73] diet makes further steps to reduce animal food intakes as vegetarians don't eat red meat, poultry and fish. It should be stressed, however, that there are several categories of vegetarians (lactovegetarians, lacto-ovo-vegetarians, etc.) and some of them use dairy products and eggs in their diet. Therefore, the definition of the Vegetarian diet in Table 14 essentially represents a combined set of food groups and respective food intakes which are acceptable from the vegetarian point of view. The data in Table 14 are compiled bearing in mind the possibility of a properly balanced diet so that all nutritional needs in terms of vitamin and mineral intake could be satisfied. It is worth noting that a driving factor of vegetarianism for many people may not be directly related to health considerations. Instead, reasons for following the Vegetarian diet might be linked to religion, morality, ethics or protection of environment [13]. Therefore, the Vegetarian diet is less restrictive than that of vegans who completely exclude all animal-derived food items from their diet. The Vegan diet is briefly discussed in the next section.

3.10.4 Vegan Diet

As already mentioned, the Vegan diet completely eliminates meat, fish and poultry. In addition to that eggs, dairy products, and honey are also excluded [13, 74]. As a result, their diet may be potentially deprived of protein, vitamin B12, calcium, iron and zinc [13, 75, 76], which are present in significant amounts in red meat and dairy products. In order to compensate for some vitamins and minerals found mainly in food of animal origin, vegans follow the concept of eating a wide variety of foods. They believe that each

legume, vegetable or fruit (and especially dark leafy vegetables and strongly colored fruits) has its own unique nutritional qualities. Thus, vegans try to eat about 100 g of leafy vegetables, like kale, lettuce, arugula, and chard, on a daily basis to provide adequate calcium intake. Protein in the vegan diet is supplied by legumes like lentils, peas and beans, as well as soy milk. Iron and zinc are found in sufficient amounts in nuts, seeds, and whole grains. The only vitamin which cannot be reliably derived from vegetables, grains and fruits is vitamin B12. To prevent the lack of vitamin B12, vegans are advised to take B12 supplements. The food intake patterns generated in this project don't include any vitamin or mineral supplements and, therefore, the vegan diet patterns obtained in this work may provide B12-vitamin intake below recommended levels.

Another important aspect of vegan diet is that it follows the latest guidelines related to healthy eating. Thus, vegans generally avoid eating processed and fast foods like white bread, cookies, crackers and refined sugar. They also completely eliminate food items containing trans fats, like margarine, donuts, and biscuits, from their diet. The vegan position on alcohol consumption is characterized by conditional acceptance. In particular, those wines which have been fined without animal-derived substances (like blood, bone marrow etc.) and beers which have been filtered by non-animal filters are acceptable and can be consumed. All these considerations were taken into account while compiling Table 14 and generating vegan food intake patterns. The alcoholic beverages are included into the vegan diet under the assumption that they satisfy the aforesaid requirements.

3.11 Reverse Risk Assessment Results

The generated database of food intake patterns needs to be validated against major macro-parameters such as an average risk of contracting a cardiovascular disease for each diet. This information is also helpful in estimating the healthy effect of more-restricted diets like Vegan or Mediterranean. Another interesting question is related to comparison of Vegetarian versus Mediterranean diet and their effect on reducing the disease risk as compared to the Traditional American diet. To answer these and other questions there have been randomly generated 4 sets of food intake patterns for each of the diets listed in Table 14. Each set comprised of 200 patterns so that the total number of patterns was 800. The pattern food items have been selected in accordance with the diet type definitions from Table 14.

Each pattern set produced a distribution of patterns over the cardiovascular disease (CVD) risk and those distributions are shown in Fig. 60 for the Vegan and Mediterranean diets and in Fig. 61 for the Vegetarian and Traditional American diets. There are several striking features which are worth mentioning. First, all the distributions are characterized by a quasi-Gaussian peak accompanied by a tail stretching into the region of the higher probability contracting a cardiovascular disease. Secondly, only few patterns fall into the low-risk range of 0 – 5%. However, the Vegan pattern distribution is characterized by a sharp peak in vicinity of 12-13% and relatively small tail on the right-side of the cupola-shape peak. On the contrary, the Traditional American diet pattern distribution exhibits a much broader peak with a central point at about 35%. The

high-risk tail of the Traditional American diet pattern distribution stretches into the region of a very high probability for contracting a disease.

Analysis of food intake details shows that this behavior is associated with availability of food products which are loaded with risk factors like sodium and saturated fat. The much more-restricted diet like Vegan essentially doesn't allow intake of these products and, thereby, greatly reduces the risk of cardiovascular diseases.

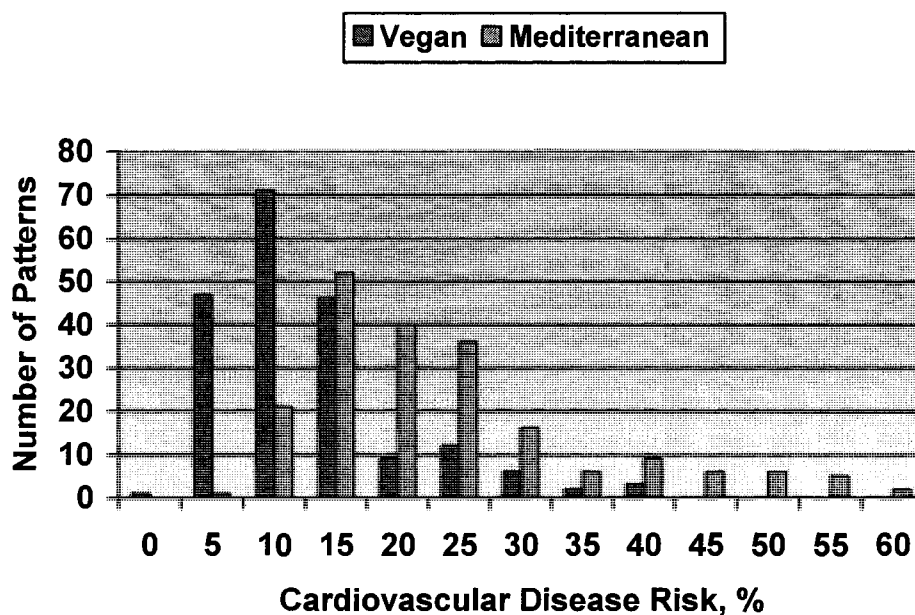


Figure 60: Risk of Cardiovascular Disease and Distribution of Vegan and Mediterranean Diet Patterns.

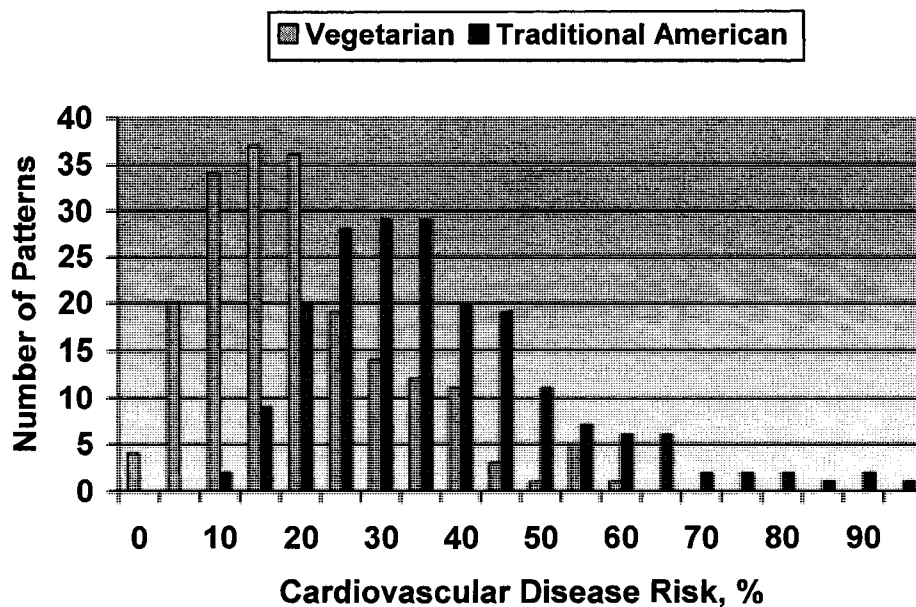


Figure 61: Risk of Cardiovascular Disease and Distribution of Vegetarian and Traditional American Diet Patterns.

The pattern distributions for Vegetarian and Mediterranean diets look in many ways similar. In particular, both distributions reach a maximum at the vicinity of the risk value 18%, which is higher than that of Vegan (12-13%), but is still much lower as compared to the American Traditional diet (35%). Only about 5% of Vegetarian diet patterns and 6% of the Mediterranean diet patterns fall into the risk range of 41% or higher. For comparison, about 40% of patterns are characterized by a risk of 41% or higher in the case of the Traditional American diet. However, despite similarities the Vegetarian diet pattern distribution has a wider left-hand side shoulder in the important region of low risk values. Thus, about 12% of the Vegetarian patterns fall into the range of 0-10% of cardiovascular disease risk versus 0.5% of the Mediterranean diet patterns.

The pattern distributions in Figs. 60 and 61 have been used for calculating the average risk of contracting stroke, heart disease and hypertension due to the diet-related risk factors. The results of these calculations are displayed in Fig. 62 for each diet type considered. Thus, Fig. 62 displays averages obtained by integrating the areas beneath the distributions in Figs.60 and 61 and then dividing the results by the total number of patterns in a set. The total risk of contracting a cardiovascular disease is also shown for reference purposes. As it follows from Fig.62, the Vegan diet produces the lowest probability of contracting a disease, whereas the Traditional American diet – the highest. Although, no detailed analysis was made of how adequate the patterns are in terms of nutrient intake, a quick examination of the pattern sets revealed that the majority of nutrients in the patterns were close to or within the recommended intake ranges.

It is interesting to compare data obtained from Monte Carlo pattern simulation with the prevalence of cardiovascular diseases in this country for people aged 65 or higher. It should be stressed, however, that data in Fig. 62 describe the probability to contract diseases at the end of the lifespan for those people who are not obese, physically active and don't smoke. The relative number of such people in the overall US population is about 20% and, therefore, is relatively small. The influence of non-diet risk factors on the disease risk can be accounted for by the relative risk coefficients as described in Chapter 2. To obtain the total probability to a contract a disease given all diet and non-diet risk factors involved one needs to multiply purely diet-related probability by the factor 1.7. Table 15 shows the results of such calculations for the heart disease and hypertension, which are compared to the statistical data available in literature [77, 78].

Average Diet-Related Risk of Cardiovascular Diseases

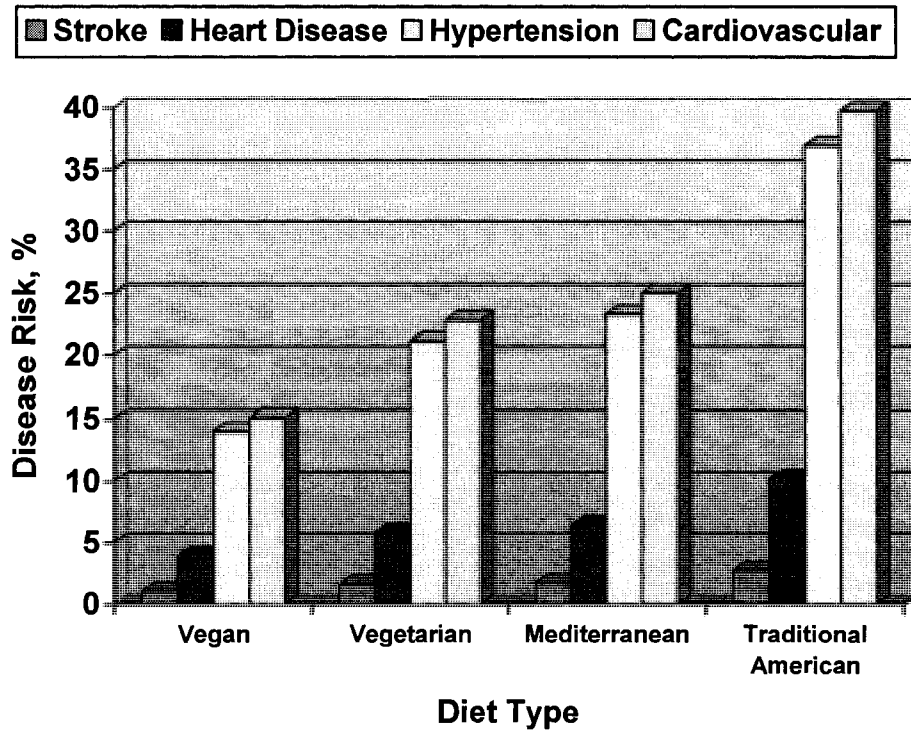


Figure 62: Average Diet-Related Risk of Contracting a Cardiovascular Disease.

Table 15: Prevalence of Heart Disease and Hypertension: Comparison of Results Obtained with Statistical Data.

Disease	Prevalence, %		
	Reference [77]	Reference [78]	Present Work
Heart Disease	19.6	n/a	17.3
Hypertension	n/a	70.8	62.7

As it follows from Table 15, the probability to contract heart disease calculated from pattern generation results is 17.3% which is somewhat lower than the statistical value 19.6%. Similarly, the risk of hypertension obtained in this work 62.7% is smaller than that from [78]. These differences are expected and can be attributed to nonlinear effects of one disease affecting the development of another in old adults. The present model is based on the linear approximation and, therefore, doesn't account for such effects. Another reason for this discrepancy might be related to alcohol consumption which has not been accounted here. Although moderate alcohol consumption tends to decrease the risk of heart disease, excessive alcohol intake may have a significant adverse effect on the human being health, including cardiovascular system. The pure alcohol consumption in USA is around 7-8 liters per capita per annum, which is at the high end of the recommended moderate alcohol intake. Excluding children, the alcohol consumption by adults is even higher. That means, on average adults consume more alcohol than what is considered an acceptable alcohol intake range. As a result, the present model predicts slightly lower, by about 10%, disease probabilities. Nonetheless, the risk values found in this work are in good quantitative agreement with the statistical data [77, 78]. The good correlation of statistical and pattern generation data on the probabilities of contracting heart disease and hypertension serves as an additional substantiation of the assumptions underlying the mathematical model developed in Chapter 2.

3.12 Conclusion

The work described in this thesis had been started with several objectives in mind. Those included (a) analysis of diet-related risk factors and their role in chronic diseases, (b) development of a mathematical model which relates the risk factors with the probability of contracting a disease, (c) building a reverse risk assessment system which would enable the user to quickly access various scenarios associated with a certain degree of risk, (d) applying JavaServer Faces technology and open source components for developing a web based application. However, most importantly the goal of this work was to demonstrate how a combination of the reverse risk assessment concept and the power of software with efficient search and scenario generation capabilities can produce a solution to a problem which otherwise would be difficult, if feasible at all, to achieve.

In this respect, it should be emphasized that any reverse risk assessment issue of practical interest involves a good deal of complexity and many parameters to be accounted for. The analysis of diet-related risk factors conducted in Chapter 1 created conditions for articulating a problem of estimating the diet-related probability of contracting a chronic disease. The problem was then mathematically formulated in Chapter 2 and a number of important equations were derived. The mathematical formalism developed in Chapter 2 was applied to cardiovascular diseases and basic assumption were verified against numerous statistical data.

Although being important, the mathematical formalism itself represented just a part of the solution to the reverse risk assessment problem. The model was supplemented by an efficient pattern generation algorithm based on the Monte Carlo approach of

randomly selecting food items from the food composition database. In addition to that, a number of other challenging issues like overall system architecture, presentation layer and pattern database design were successfully resolved in the project.

Usage of JavaServer Faces technology and Java Sun Studio Creator IDE during the software implementation phase contributed greatly a relatively fast completion of the reverse risk assessment application, as described in Chapter 3. Combining JSF technology with the open source components (Apache Tomcat application server and MySQL database management system) resulted in building a portable and scalable web-based application. The reverse risk assessment system built in course of the project is capable of producing realistic estimates for the risk of contracting a cardiovascular disease for various diet types. The computational results are in good agreement with the statistical data on the prevalence of hypertension and heart disease available in literature. Since data obtained by Monte Carlo simulation schemes are often considered as theoretical experiments, the system can be regarded as a kind of experimental setup which can be used for further research work.

There is no doubt that the results of pattern generation and their implications discussed in Chapter 3 constitute a small portion of what can be produced by the system. In particular, it would be interesting to analyze food intake patterns on the subject of nutrient deficiency in more detail and make quantitative recommendations for improving nutrient intake for certain groups of people like vegans and vegetarians. Another exciting area for further research might be identification of the so-called “ideal diet” – a set of

food intake patterns with an adequate nutrient intake and a minimal risk of cardiovascular disease. These research areas are, however, beyond the scope of the present project.

Appendix A. Use Case Sequence Diagrams

A.1 Use Case: Login

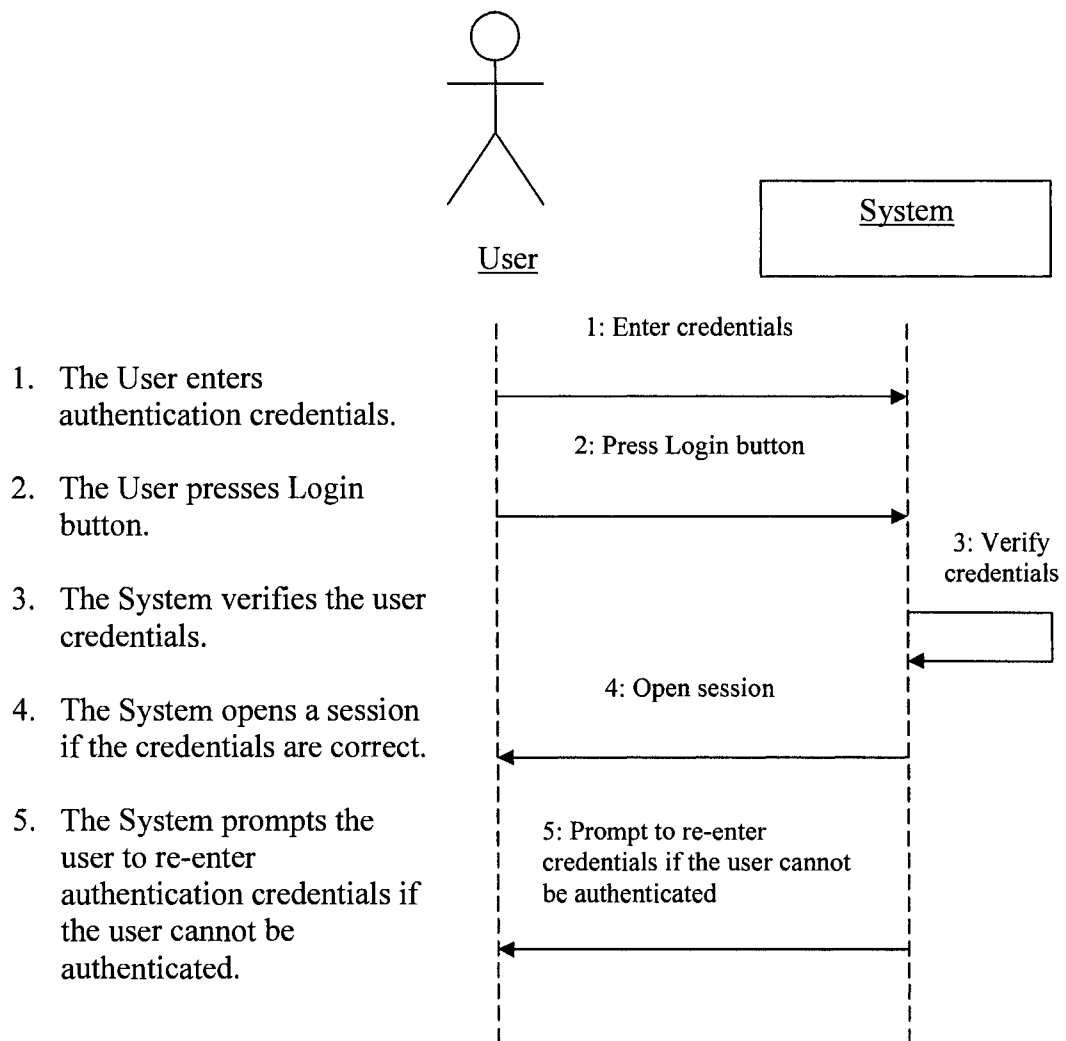


Figure 63: Login Use Case Sequence Diagram.

A.2 Use Case: Update Profile

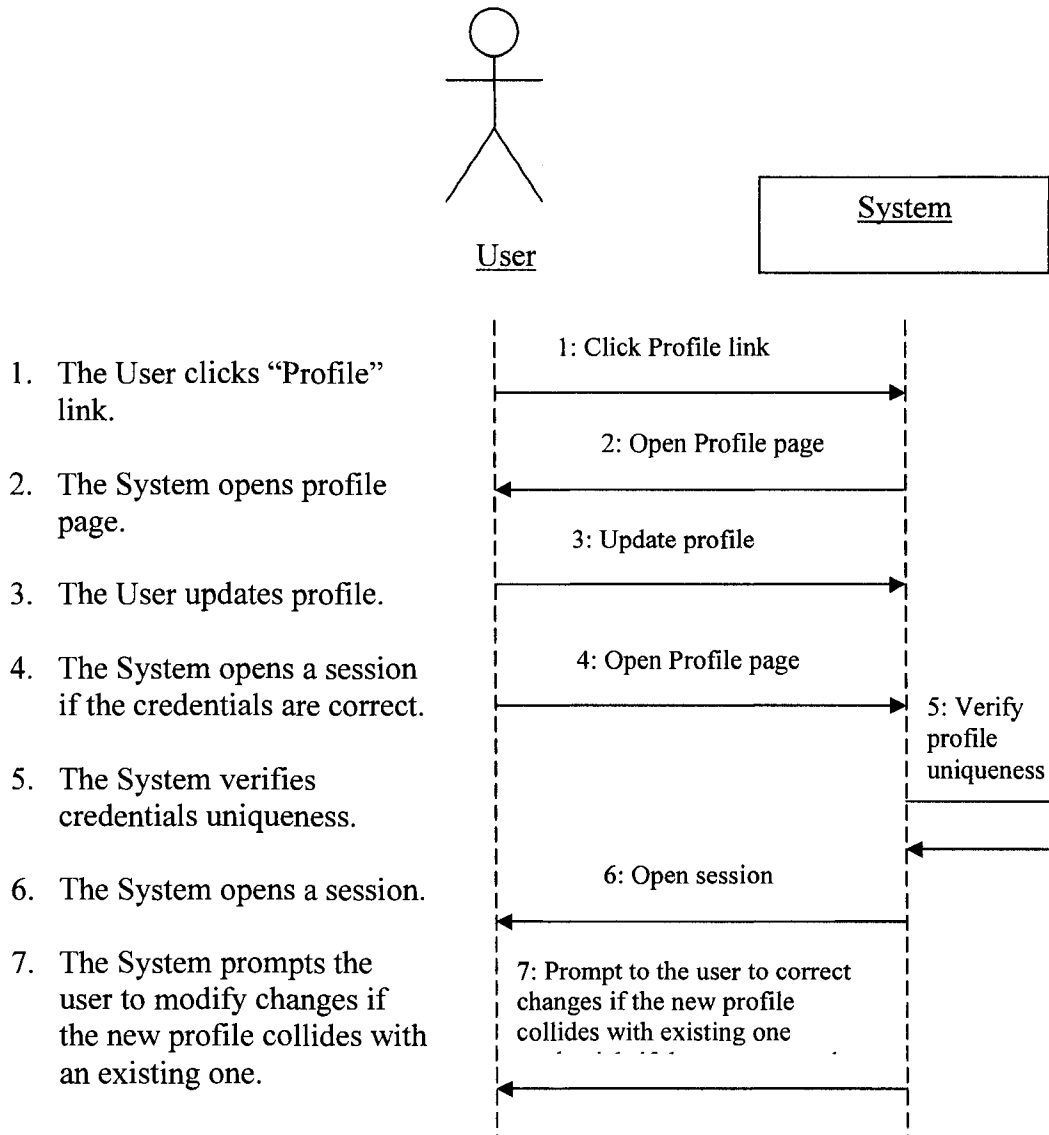


Figure 64: Update Profile Use Case Sequence Diagram.

A.3 Use Case: Look up Calendar

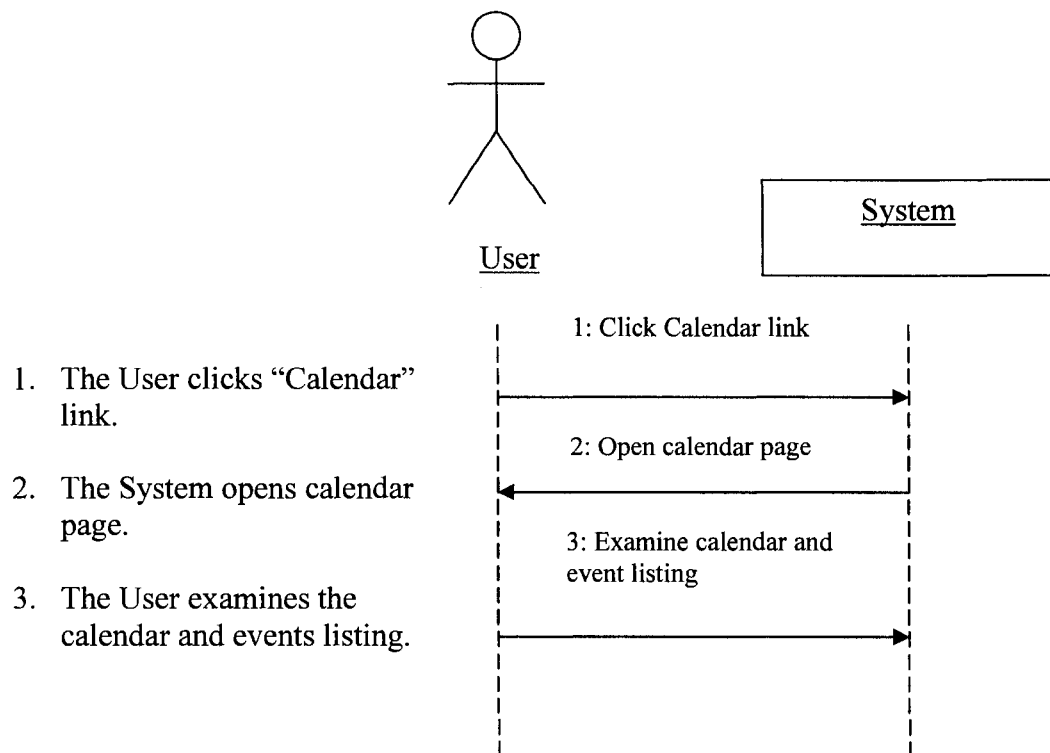


Figure 65: Look up Glossary Use Case Sequence Diagram.

A.4 Use Case: Calculate Risk

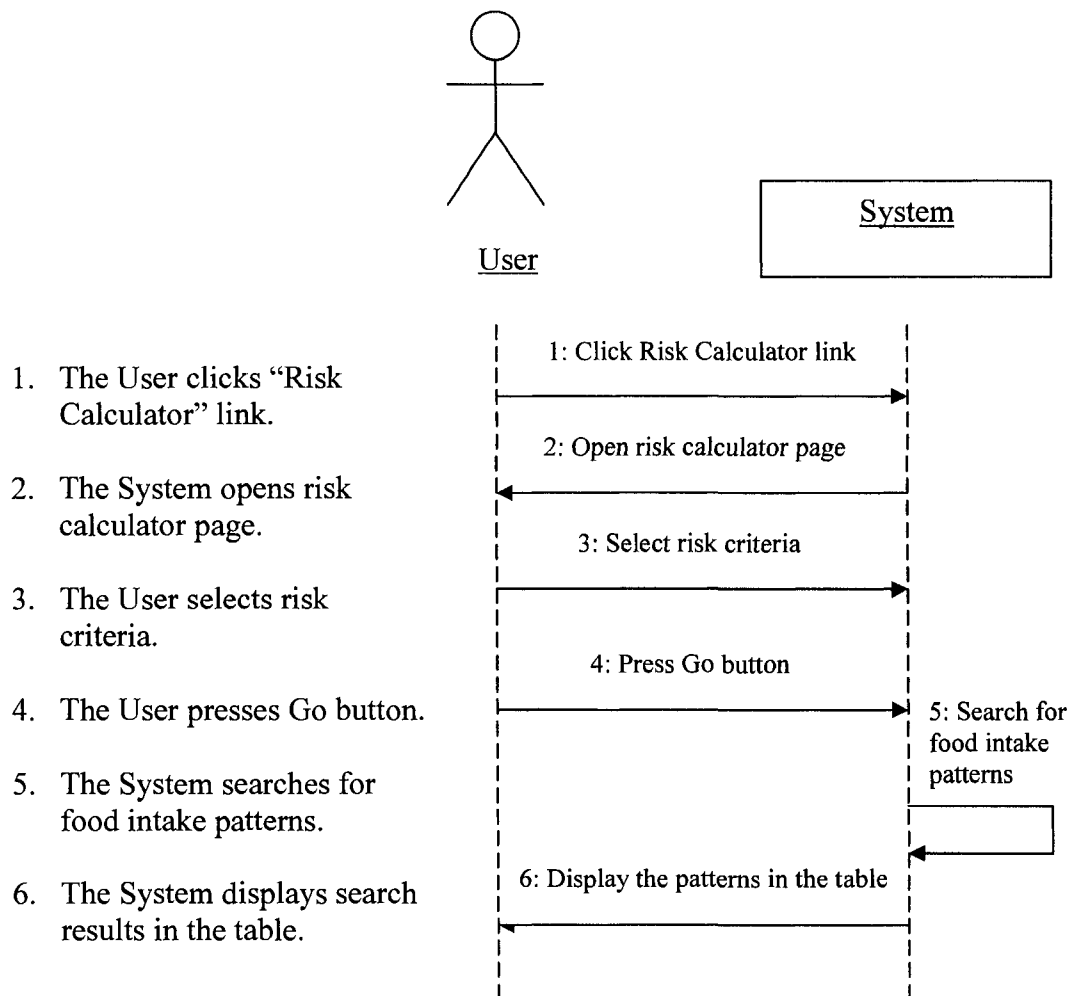


Figure 66: Calculate Risk Use Case Sequence Diagram.

A.5 Use Case: Browse Results

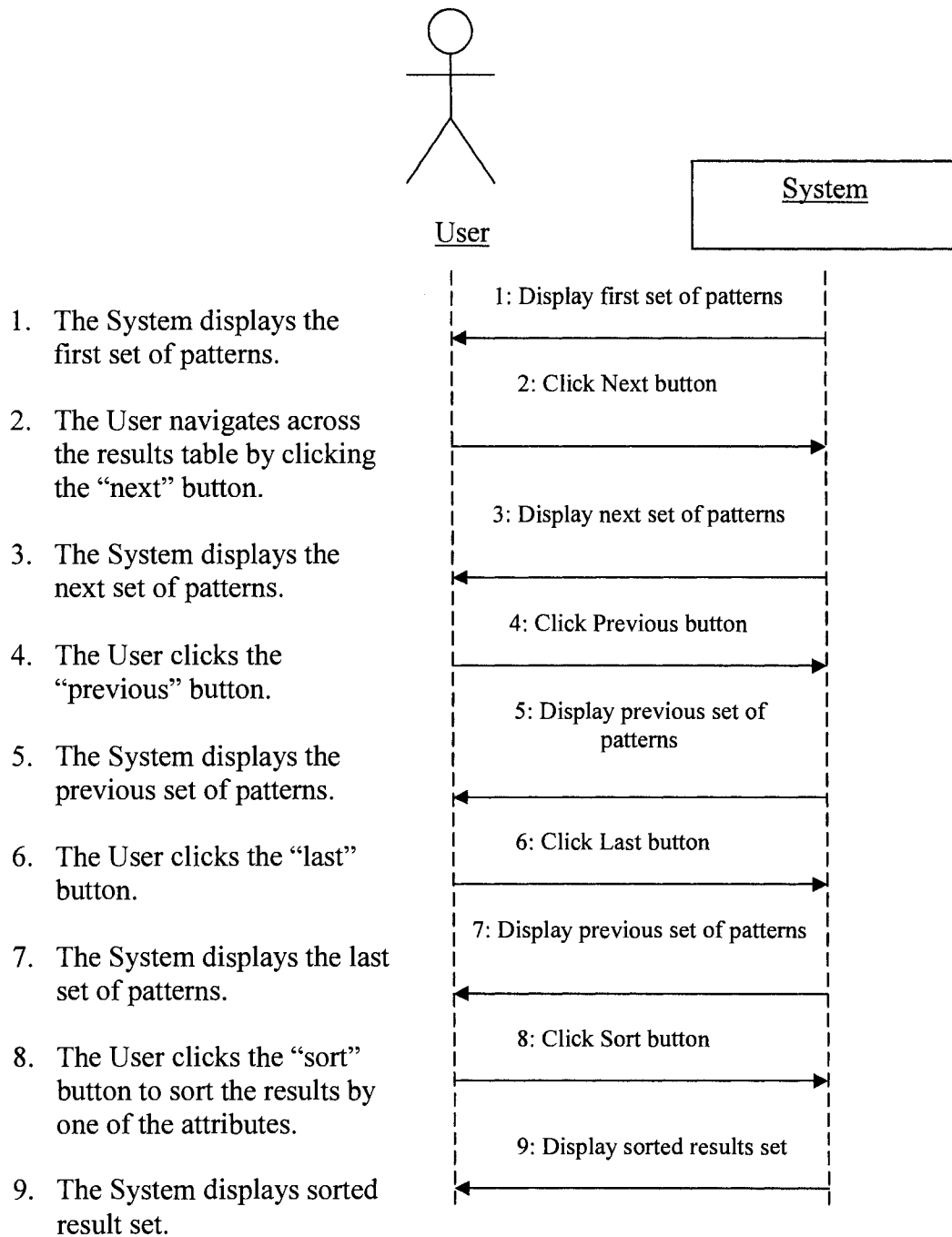


Figure 67: Browse Results Use Case Sequence Diagram.

A.6 Use Case: Explore Pattern

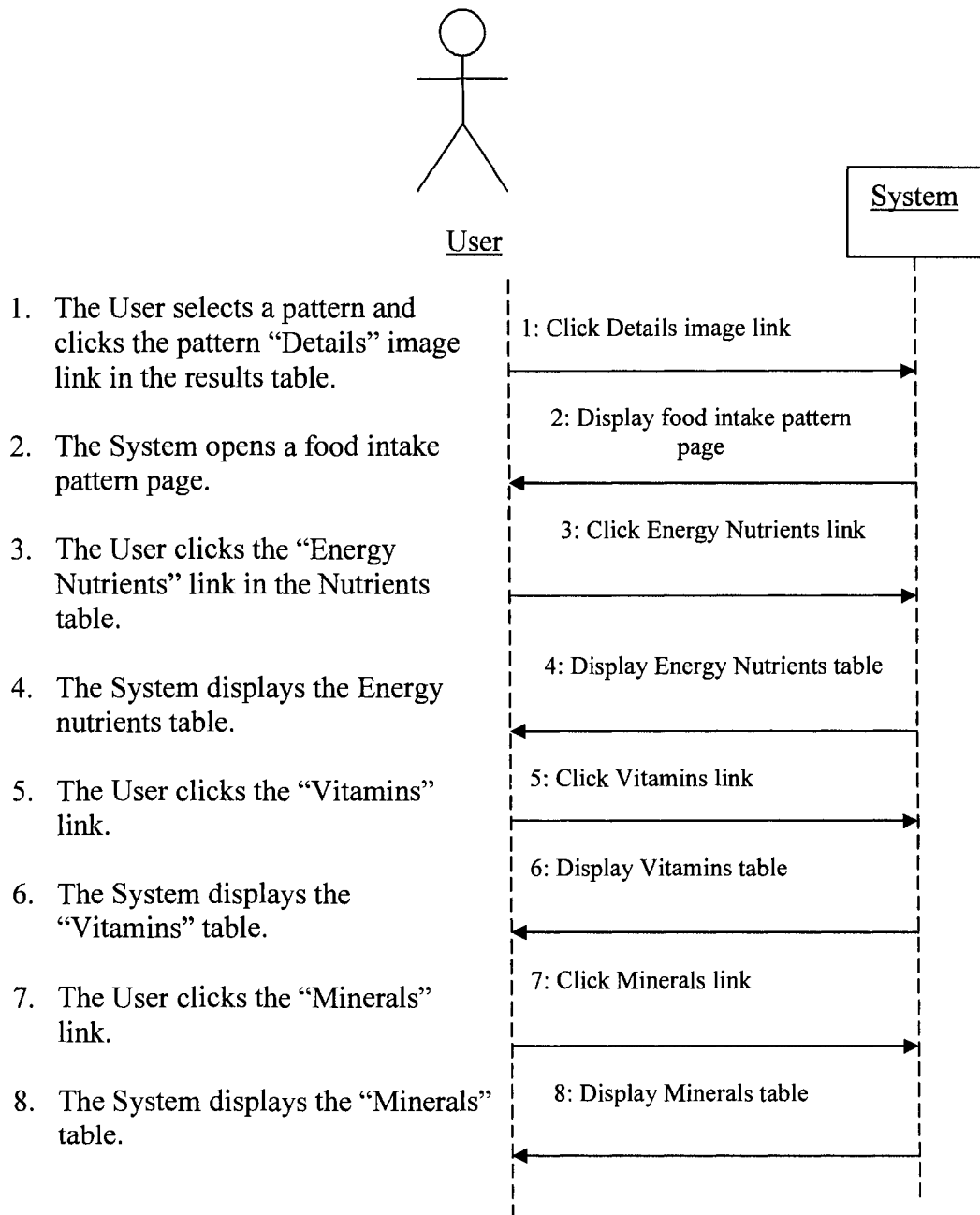


Figure 68: Explore Pattern Use Case Sequence Diagram.

A.7 Use Case: Look up Glossary

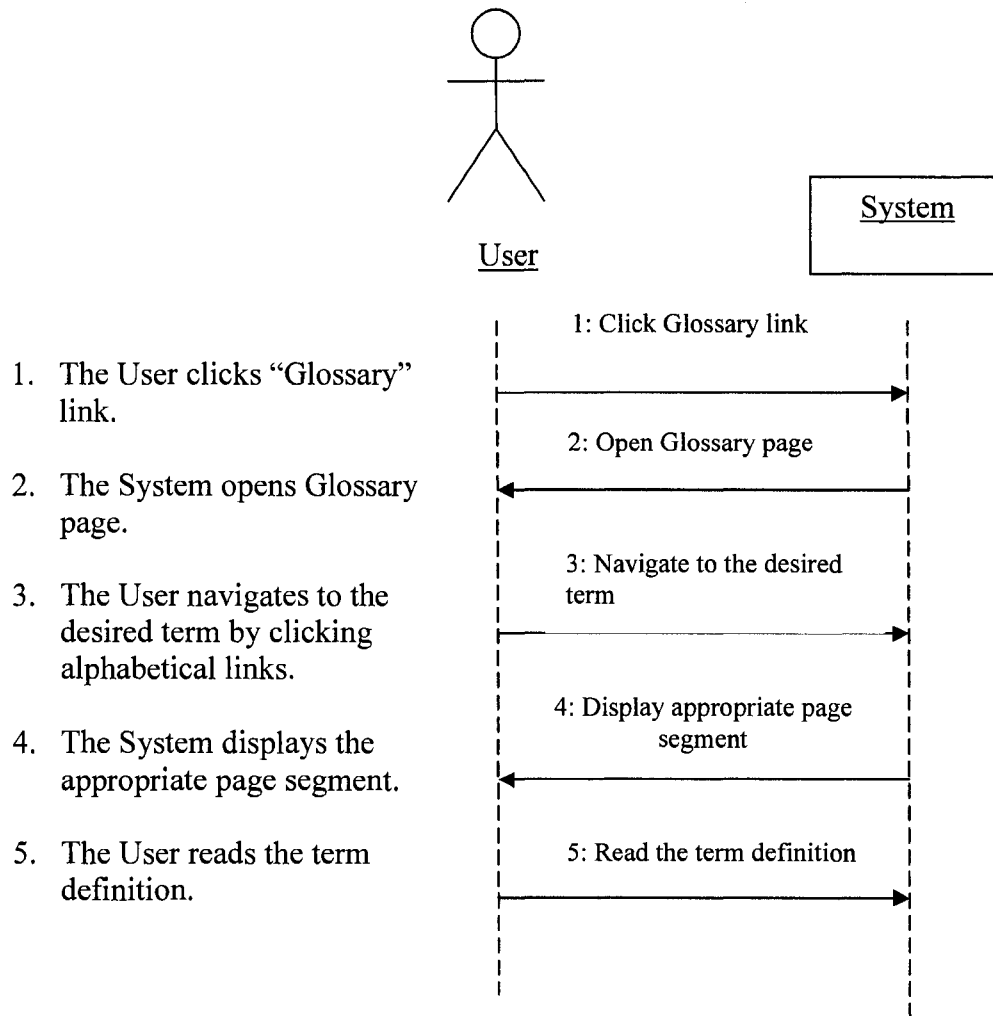


Figure 69: Look up Glossary Use Case Sequence Diagram.

A.8 Use Case: Read General Information

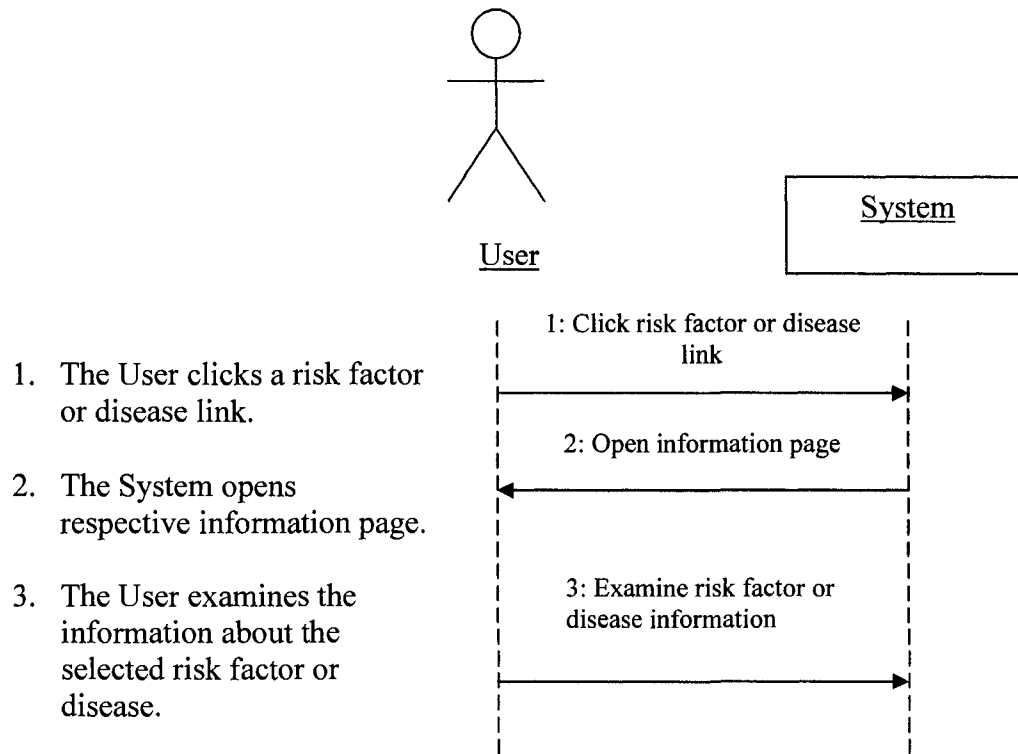


Figure 70: Read General Info Use Case Sequence Diagram.

A.9 Use Case: Use External Link

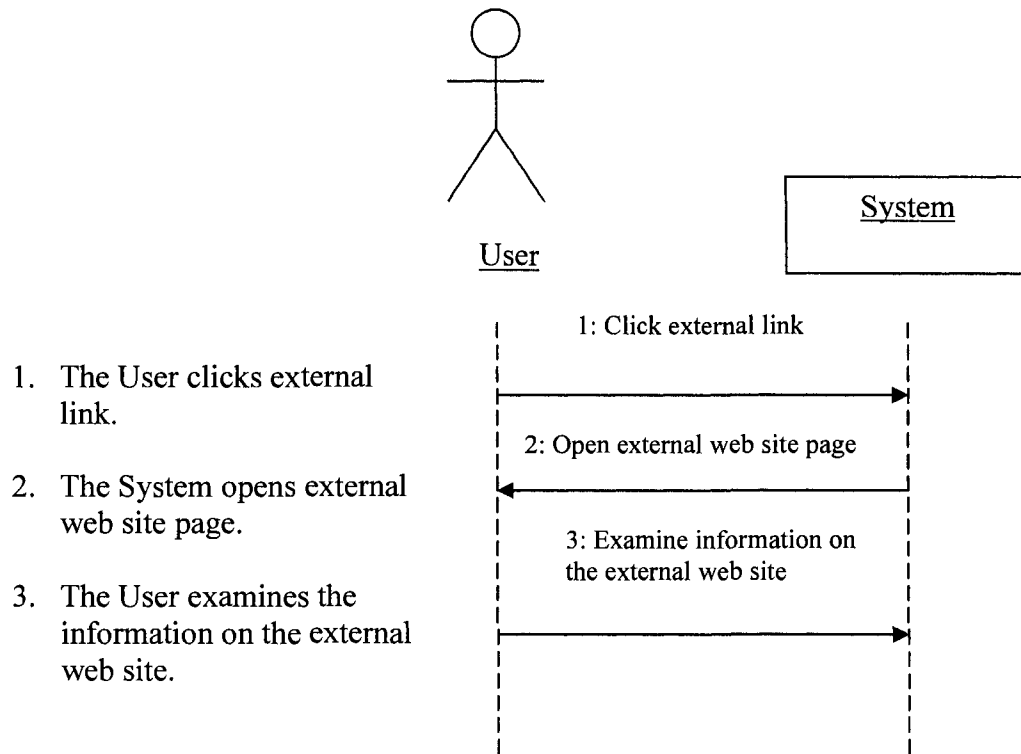


Figure 71: Use External Link Use Case Sequence Diagram.

A.10 Use Case: Create New User Profile

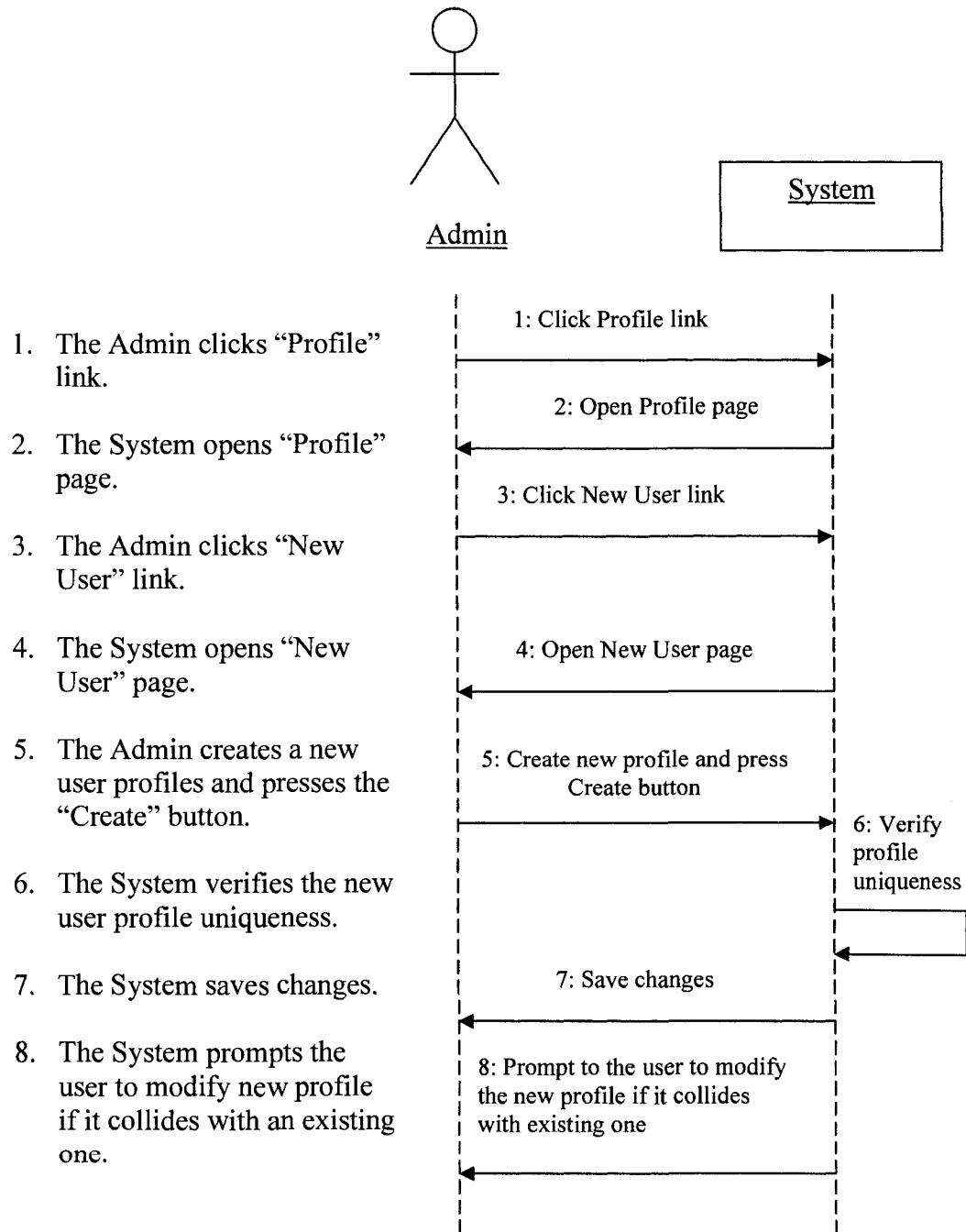


Figure 72: Create New User Profile Use Case Sequence Diagram.

Appendix B. Food Intake Pattern DTD

This appendix contains the FoodIntakePattern.dtd code which serves as document type definition (DTD) for food intake pattern description provided by a respective XML file. The elements “Quantity” and “TotalEnergy” pertain to the total amount of food and calorie intake in the pattern. The element “Person” is reserved for future use. The remaining elements definition is pretty self-descriptive.

```
<!-- Author: Igor Tilinin -->
<!-- Filename: FoodIntakePattern.dtd -->
<!-- Project: Reverse Risk Assessment -->
<!--FoodINtakePattern -->
<!ELEMENT FoodIntakePattern (Person, Quantity, TotalEnergy, Diseases, Nutrients,
FoodGroup+ ) >
<!-- Person -->
<!ELEMENT Person (Age, Gender, Ethnicity, Weight, BMI) >
<!ELEMENT Age (#PCDATA) >
<!ELEMENT Gender (#PCDATA) >
<!ELEMENT Ethnicity (#PCDATA) >
<!ELEMENT Weight (#PCDATA) >
<!ELEMENT BMI (#PCDATA) >
<!ELEMENT Quantity (#PCDATA) >
```

```
<!ELEMENT TotalEnergy (#PCDATA) >
<!-- Diseases -->
<!ELEMENT Diseases (Disease)* >
<!ELEMENT Disease (Description, Risk) >
<!ELEMENT Description (#PCDATA) >
<!ELEMENT Risk (#PCDATA) >
<!-- Nutrients -->
<!ELEMENT Nutrients (NutrientGroup)+ >
<!ELEMENT NutrientGroup (Nutrient)+ >
<!ELEMENT Nutrient (NutrientDescription, AmountAvailable, AmountRecommended,
LowerLimit, UpperLimit, MeasurementUnit) >
<!ELEMENT NutrientDescription (#PCDATA) >
<!ELEMENT AmountAvailable (#PCDATA) >
<!ELEMENT AmountRecommended (#PCDATA) >
<!ELEMENT UpperLimit (#PCDATA) >
<!ELEMENT LowerLimit (#PCDATA) >
<!-- FoodGroup -->
<!ELEMENT FoodGroup (Composition, FoodItem*) >
<!ELEMENT Composition (#PCDATA) >
<!ELEMENT FoodItem (Amount, FoodItemQuantity, Energy, NutrientItem+) >
<!ELEMENT NutrientItem (NutrientItemQuantity, MeasurementUnit) >
<!ELEMENT Amount (#PCDATA) >
```

<!ELEMENT FoodItemQuantity (#PCDATA) >
<!ELEMENT Energy (#PCDATA) >
<!ELEMENT NutrientItemQuantity (#PCDATA) >
<!ELEMENT MeasurementUnit (#PCDATA) >
<!-- Attribute List -->
<!ATTLIST FoodIntakePattern patternID ID #REQUIRED>
<!ATTLIST FoodIntakePattern type (Vegan|Vegetarian|Mediterranean|Traditional)
#REQUIRED>
<!ATTLIST FoodIntakePattern frequency (Meal|Daily|Weekly|Monthly) #REQUIRED>
<!ATTLIST Diseases type CDATA #REQUIRED >
<!ATTLIST Disease name CDATA #REQUIRED >
<!ATTLIST NutrientGroup name CDATA #REQUIRED >
<!ATTLIST Nutrient name CDATA #REQUIRED >
<!ATTLIST FoodGroup name CDATA #REQUIRED >
<!ATTLIST FoodItem name CDATA #REQUIRED >
<!ATTLIST NutrientItem name CDATA #REQUIRED >
<!ATTLIST NutrientItem group CDATA #REQUIRED >
<!ATTLIST NutrientItem quantity CDATA #REQUIRED >

Appendix C. XSL Stylesheet Code

The FoodIntakePattern.xsl stylesheet plays an important role in transforming the pattern XML code into HTML format when being invoked from the browser. Below is the full version of the XSL code.

```
<?xml version="1.0"?>
<!-- Author: Igor Tilinin -->
<!-- Filename: FoodIntakePattern.xml - template -->
<!-- Project: Reverse Risk Assessment -->
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:output method="html" version="4.0"/>
<xsl:param name="energyYielding">Energy-Yielding Nutrients</xsl:param>
<xsl:param name="nutrientGroup"></xsl:param>
<xsl:template match="/">
<html>
<head>
<title>Food Intake Pattern</title>
<link href="FoodIntakePattern.css" rel="stylesheet" type="text/css" />
</head>
<body>
```

```

<H1 class="title"><a name="top">Food Intake Pattern</a></H1>
<div style="clear:left; width: 640; margin-top: 20">
<table border="0" class="general">
<tr>
  <td width="30%" valign="top">
    <table ><tr><td><table border="1"><tr><td>
      <table border="0" class="diseases">
        <tr>
          <th align="left" class="head1"
            colspan="2">Disease Risk, %</th>
        </tr>
        <xsl:apply-templates
          select="FoodIntakePattern/Diseases/Disease">
        </xsl:apply-templates>
      </table>
    </td></tr></table>
  </td>
</tr>
<tr>
  <td class="spacer"><font color="white"></font></td>
</tr>
<tr>

```

```

<td>
  <table border="1"><tr><td>
    <table border="0" class="nutrients">
      <tr>
        <th align="left" class="head1">Nutrients</th>
      </tr>
      <xsl:apply-templates
select="FoodIntakePattern/Nutrients/NutrientGroup/@name">
        </xsl:apply-templates>
      </table>
    </td></tr></table>
  </td>
</tr>
</table>
</td>
<!-- Pattern Composition Table -->
<td width="70%"><table border="1"><tr><td>
  <table border="0" class="food-intake-composition">
    <tr>
      <th align="left" class="head1" colspan="2">Daily Food Intake
    </th>
    </tr>
  </table>
  </td>

```

```

        <xsl:apply-templates select="FoodIntakePattern/FoodGroup">
            </xsl:apply-templates>
        </table>
    </td></tr></table>
</td>
</tr>
</table>
</div>
<!-- Display 5 tables with details of nutrient availability in the food items -->
<xsl:apply-templates select="FoodIntakePattern/Nutrients/NutrientGroup">
</xsl:apply-templates>
</body>
</html>
</xsl:template>
<!-- Templates Start Here -->
<!-- Disease -->
<xsl:template match="Diseases/Disease">
<tr><td class="cell1" width="70%">
        <xsl:value-of select="./@name"/>
    </td>
    <td align="center" width="30%" class="cell3" >
        <xsl:value-of select="./Risk"/>
    </td>
</tr>

```

```

        </td>
</tr>
</xsl:template>
<!-- NutrientGroup -->
<xsl:template match="Nutrients/NutrientGroup/@name">
<tr><td class="cell1"><xsl:element name="a">
        <xsl:attribute name="href"##<xsl:value-of select="."/ />
        </xsl:attribute>
        <xsl:attribute name="color">#989898
        </xsl:attribute>
        <xsl:value-of select="."/ />
        </xsl:element>
    </td>
</tr>
</xsl:template>
<!-- Food Group Template -->
<xsl:template match="FoodGroup">
<xsl:variable name="varValue">Empty</xsl:variable>
<xsl:if test="./Composition != $varValue">
<tr>
<xsl:if test="position() mod 2 = 0">
        <xsl:attribute name="bgcolor">#DDEEDD</xsl:attribute>

```



```

</xsl:if>

<td class="cell2" valign="top" width="30%">

    <xsl:value-of select="./@name"/>

</td>

<td class="cell3" width="70%">

    <xsl:value-of select="./Composition"/>

</td>

</tr>

</xsl:if>

</xsl:template>

<!-- Nutrient Group template -->

<xsl:template match="Nutrients/NutrientGroup">

<div style="clear:left; width: 640; margin-top: 20">

<br/><center>

<table border="1" width="900" >

    <tr><td><table class="nutrient-composition" width="900">

    <tr><th colspan="11" class="head2">

        <xsl:element name="a">

            <xsl:attribute name="name"><xsl:value-of select="./@name"/>

            </xsl:attribute>

            <xsl:value-of select="./@name"/>

            </xsl:element>

```

```

        </th>
    </tr>
<tr>
    <th width="15%" class="head3">Food Item</th>
    <th width="8%" class="head3">Quantity</th>
    <xsl:if test="$energyYielding = ./@name">
        <th width="8%" align="center" class="head3">Energy, kcal</th>
    </xsl:if>
    <!-- Display nutrient names in the table header -->
        <xsl:apply-templates select="./Nutrient/@name">
        </xsl:apply-templates>
    </tr>
    <!-- Actual Content -->
        <xsl:apply-templates select="../../FoodGroup/FoodItem">
            <xsl:with-param name = "nutrientGroup" select="./@name"/>
        </xsl:apply-templates>
    <!-- Totals -->
<tr><td class="cell6" colspan="2">Total:</td>
    <xsl:if test="$energyYielding = ./@name">
        <td class="cell6" >
            <xsl:apply-templates select="../../TotalEnergy">
            </xsl:apply-templates></td>
    </xsl:if>

```

```

    </xsl:if>

    <xsl:apply-templates select="./Nutrient/AmountAvailable">
    </xsl:apply-templates>
</tr>

<!-- Energy Nutrients RDA -->

<tr>

<xsl:attribute name="bgcolor">#DDFFDD</xsl:attribute>

<xsl:if test="$energyYielding = ./@name">

    <td class="cell7" colspan="3">Recommended Daily Allowance (RDA)</td>

</xsl:if>

<xsl:if test="$energyYielding != ./@name">

    <td class="cell7" colspan="2">Recommended Daily Allowance (RDA)</td>

</xsl:if>

    <xsl:apply-templates select="./Nutrient/AmountRecommended">
    </xsl:apply-templates>

</tr>

<tr>

<xsl:attribute name="bgcolor">#DDFFDD</xsl:attribute>

<xsl:if test="$energyYielding = ./@name">

    <td class="cell7" colspan="3">Tolerable Range</td>

</xsl:if>

<xsl:if test="$energyYielding != ./@name">

```

```

        <td class="cell7" colspan="2">Tolerable Range</td>
</xsl:if>
        <xsl:apply-templates select="./Nutrient/LowerLimit">
        </xsl:apply-templates>
</tr>
</table>
</td></tr>
</table>
</center>
</div>
<table><tr><td class="cell7">
<a href="#top">Back to Top</a>
</td></tr>
</table>
</xsl:template>
<xsl:template match="Nutrients/NutrientGroup/Nutrient/AmountAvailable">
        <td class="cell6"><xsl:value-of select="."/></td>
</xsl:template>
<xsl:template match="Nutrients/NutrientGroup/Nutrient/AmountRecommended">
        <td class="cell7"><xsl:value-of select="."/></td>
</xsl:template>
<xsl:template match="Nutrients/NutrientGroup/Nutrient/LowerLimit">

```

```

        <td class="cell7"><xsl:value-of select="."/> - <xsl:value-of
select="../UpperLimit"/></td>
</xsl:template>
<xsl:template match="FoodIntakePattern/Quantity">
    <td><xsl:value-of select="."/></td>
</xsl:template>
<xsl:template match="Nutrients/NutrientGroup/Nutrient/@name">
    <th width="10%" align="center" class="head3"><xsl:value-of select="."/>,
<xsl:value-of select="../MeasurementUnit"/></th>
</xsl:template>
<!-- Template FoodItem -->
<xsl:template name="foodItem" match="FoodIntakePattern/FoodGroup/FoodItem">
<xsl:param name="nutrientGroup"/>
<tr>
<xsl:if test="position() mod 2 = 0">
    <xsl:attribute name="bgcolor">#DDFFDD</xsl:attribute>
</xsl:if>
    <td align="left"><xsl:value-of select="./@name"/></td>
    <td align="left"><xsl:value-of select="./Amount"/></td>
<xsl:if test="$energyYielding = $nutrientGroup">
    <td align="left"><xsl:value-of select="./Energy"/></td>
</xsl:if>

```

```

        <xsl:apply-templates select="./NutrientItem">
            <xsl:with-param name = "nutrientGroup" select="$nutrientGroup"/>
        </xsl:apply-templates>
    </tr>
</xsl:template>
<xsl:template match="FoodIntakePattern/FoodGroup/FoodItem/NutrientItem">
    <xsl:param name="nutrientGroup"/>
    <xsl:if test="$nutrientGroup = ./@group">
        <td align="left"><xsl:value-of select="./@quantity"/></td>
    </xsl:if>
</xsl:template>
<xsl:template match="FoodIntakePattern/TotalEnergy">
    <xsl:value-of select="."/>
</xsl:template>
</xsl:stylesheet>

```

References

- [1] N.A. Doherty, *Integrated Risk Management*. New York: McGraw-Hill, 2000.
- [2] M. Crouhy, *Risk Management*. New York: McGraw-Hill, 2000.
- [3] M. Wideman, Ed., *Project and Program Risk Management, A Guide to Managing Project Risks and Opportunities*, Newton Square, PA: Project Management Institute Publications, 1992
- [4] R.A. Pollak, "Government risk regulation" in *Challenges in Risk Assessment and Risk Management, Annals of the American Academy of Political and Social Science*, vol. 545, pp. 25-34, May 1996.
- [5] W. Leiss, "Three phases in the evolution of risk communication practice," in *Challenges in Risk Assessment and Risk Management, Annals of the American Academy of Political and Social Science*, vol. 545, pp. 85-94, May 1996.
- [6] J.V. Rodricks, "Historical perspective of risk assessment and review of steps," in *Process in Food Safety Policy, Science, and Risk Assessment: Strengthening the Connection: Workshop Proceedings*, Institute of Medicine (IOM), 2001, p. 25.
- [7] K.S. Delaplane, "Pesticide usage in the United States: history, benefits, risks, and trends," Athens, GA: The University of Georgia, College of Agricultural and Environmental Sciences, Bulletin 1121, p. 5, 2000.
- [8] *ADI List – Acceptable Daily Intakes for Agricultural and Veterinary Chemicals*. Woden, ACT: Australian Government, Department of Health and Aging, 2006.
- [9] H. Hristov, N. Butler, P. Painter, and D. Siegel, "Human-exposure-based screening numbers developed to aid estimation of cleanup costs for contaminated soil," California Environmental Protection Agency, Office of Environmental Health Hazard Assessment, Sacramento, CA, pp. 1-134, November 2004, January 2005 Revision.
- [10] J. Hubbard, *Human Health Risk Assessment*, Philadelphia: U.S. Environmental Protection Agency, 2006.
- [11] R.M. Harrison and R.E. Hester. Eds., *Alternatives to Animal Testing, Issues in Environmental Science and Technology*, London: RSC Publishing 2006.

- [12] *Heart Disease and Stroke Statistics – 2006 Update*, Dallas, Texas: American Heart Association and American Stroke Association, 2006 pp. 1–40.
- [13] E.M. Whitney and S.R. Rolfes, *Understanding Nutrition*, New York: West/Wadsworth, 1999.
- [14] Centers for Disease Control and Prevention, “Prevalence of Physical Activity, Including Lifestyle Activities Among Adults --- United States, 2000-2001,” *Mortality and Morbidity Weekly Report*, vol. 52 No. 32, pp. 764-769, 2003.
- [15] *Nutrition and Your Health: Dietary Guidelines for Americans. Part D Science Base*, Washington, DC: U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2005.
- [16] L.H. Kuller, “Dietary fat and chronic diseases: epidemiological overview,” *J. Am. Diet. Ass.*, vol. 97, pp. S9-S15, 1997.
- [17] J.E. Blundell and J. I. Macdiarmid, “Fat as a risk factor for overconsumption,” *J. Am. Diet. Ass.*, vol. 97, pp. S63-S69, 1997.
- [18] American Dietetic Association, “Position of the American Dietetic Association. Health implications of dietary fiber,” *J. Am. Diet. Ass.*, vol. 97 pp. 1157-1159, 1997.
- [19] H. Whelton, “Overview of the impact of changing global patterns of dental caries experience on caries clinical trials,” *J. Dent. Res.*, vol. 83C, p.p. C29-C34, 2004.
- [20] T.A. Kotchen and J. M. Kotchen, “Dietary sodium and blood pressure: interactions with other nutrients,” *Am. J. Clinical Nutrition*, vol. 65, pp.7085-7115, 1997.
- [21] A. Lichtenstien, L. Appel, M. Brands, a.o., “Diet and lifecycle recommendations revision 2006. A scientific statement from the American Heart Association Nutrition Committee,” *Circulation*, vol. 114, pp.82-96, 2006.
- [22] D. Mzarffarian, M. B. Katan, A. Ascherio, M. J. Stampfer, W. C. Willet, “Trans fatty acids and cardiovascular disease,” *N. Engl. J Med.*, vol. 354 No.15, pp. 1601–1613, 2006.
- [23] W. H. Dietz, “Childhood obesity: susceptibility, cause, and management.” *Journal of Pediatrics*, vol. 103, No. 5, pp. 676-686, 1983.
- [24] L.A. Tucker, G. T. Seljaas, and R. L. Hager, “Body fat percentage of children varies according to their diet composition,” *J. Am. Diet. Ass.*, vol. 97, pp. 981-986, 1997.

- [25] American Diabetes Association, "All about diabetes," October 27, 2006, <http://www.diabetes.org/about-diabetes.jsp>.
- [26] *National Diabetes Fact Sheet: General Information and National Estimates in the United States, 2005*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2005.
- [27] *Bone health and osteoporosis: a report of the Surgeon General*, Washington, D.C.: U.S. Dept. of Health and Human Services, Public Health Service, Office of the Surgeon General, 2004.
- [28] *8 Common Myths About Osteoporosis*, Washington, D.C., National Osteoporosis Foundation, 2006.
- [29] M.L. Corman, S.I. Allison, J.P. Kuehne, "Diverticular disease" in *Handbook of Colon and Rectal Surgery*. Hagerstown, MD, Lippincott, Williams & Wilkins; 2002, pp. 637–653.
- [30] J.E. King, Ed., "Diverticular disease" in *Mayo Clinic on Digestive Health*, Rochester, MN, Mayo Clinic, 2000, pp. 125–132.
- [31] P.W. Marcello, "Understanding diverticular disease," *Ostomy Quarterly*, vol. 39 No. 2, p. p. 56–57, 2002.
- [32] P. Puska, S. Mendis, D. Porter, "Cardiovascular disease (CVD)," World Health Organization, Geneva, Switzerland, WHO Report, pp. 1-2, 2003.
- [33] S.M. Grundy, R. Pasternak, P. Greenland, S. Smith, Jr, V. Fuster, "Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations," *Circulation*, vol. 100, pp.1481-1492, 1999.
- [34] Joint WHO/FAO Expert Consultation, "Diet, Nutrition and Prevention of Chronic Diseases," World Health Organization, Geneva, Switzerland, WHO Tech. Report Series, No. 916, pp.1-160, 2003.
- [35] L. Buerger, "Thromboangiitis obliterations; a study of vascular lesions leading to presenile spontaneous gangrene," *Am. J. Med. Sci.*, vol. 266, pp. 278-291, 1973.
- [36] S.G. Wannamethee, A.G. Shaper, A.G. Whincup, and M. Walker, "Smoking cessation and the risk of stroke in middle aged men," *J. Am. Med. Ass.*, vol. 274, pp. 155-160, 1995.

- [37] P.A. Wolf, R. B. D'Agostino, W.B. Kannel, R. Bonita, A.J. Belanger, "Cigarette smoking as a risk factor for stroke. the Farmington study," *J. Am. Med. Ass.*, vol. 120, pp. 1025-1029, 1988.
- [38] R.M. Krauss, M Winston, and B.J. Fletcher, "Obesity: impact on cardiovascular disease," *Circulation*, vol. 98, pp. 1472-1476, 1998.
- [39] H.B. Hubert, M. Feinleib, P.M. McNamara and W.P. Castelli, "Obesity as an independent risk factor for cardiovascular disease: a 26- year follow-up of participants in the Framingham heart study," *Circulation*, vol. 67, pp. 968-977, 1983.
- [40] "Preventing and Managing Global Epidemic of Obesity," World Health Organization, Geneva, Switzerland, WHO Tech. Report Series, vol. 894, p.p.1-253, 2000.
- [41] A.J. Flint and E.B. Rimm, "Commentary: obesity and cardiovascular disease risk among the young and old—is BMI the wrong benchmark?," *International Journal of Epidemiology*, vol. 35, No. 1, pp.187-189, 2006.
- [42] A. Pajak, R. Topor-Madry, A. Waskiewicz, E. Sygnowska., "Body mass index and risk of death in middle-aged men and women in Poland. Results of POL-MONICA cohort study," *Kardiol. Pol.*, vol. 62, No. 2, pp. 95-105, 2005.
- [43] W.B. Kannel, R.B. D'Agostino, and J.L. Cobb, "Effect of weight on cardiovascular disease," *Am. J. Clinical Nutrition*, vol. 63, pp. 419S-422S, 1996.
- [44] D.A. McCarron, and M.E. Reusser, "Body weight and blood pressure regulation," *Am. J. Clinical Nutrition*, vol. 63, pp. 423S-425S, 1996.
- [45] R. Walker and P. Sareli, "Coronary heart disease: outlook for Africa," *J. Royal Soc. Med.*, vol. 90, No. 1, p.p. 23–27, 1997.
- [46] Centers for Disease Control and Prevention, "Cigarette smoking among adults - United States, 2004," *Morbidity and Mortality Weekly Report*, vol. 54, No. 44, pp. 1121-1124, 2005.
- [47] N. Haapanen-Niemi, S. Miilunpalo, M. Pasanen, I. Vuori, P. Oja, and J. Malmberg, "Body mass index, physical inactivity and low level of physical fitness as determinants of all-cause and cardiovascular disease mortality - 16 y follow-up of middle-aged and elderly men and women," *International Journal of Obesity*, vol. 24, No. 11, pp. 1465-1474, 2000.

- [48] *Prevalence of Overweight and Obesity Among Adults: United States, 1999-2002*, Atlanta, GA, Centers for Disease Control and Prevention, National Center for Health Statistics, 2003.
- [49] K.M. Flegal, M.D.Carroll, C.L. Ogden, and C.L. Johnson, "Prevalence and trends in obesity among US adults, 1999-2000," *J. Am. Med. Ass.*, vol. 288, pp.1723-1727, 2002.
- [50] A.A. Hedley, C.L. Ogden, C.L. Johnson, a.o., "Overweight and obesity among US children, adolescents, and adults, 1999-2002," *J. Am. Med. Ass.*, vol. 291, pp. 2847-285, 2003.
- [51] A. Lichtenstien, L. Appel, M. Brands, a.o. "Diet and lifecycle recommendations revision 2006. A scientific statement from the American Heart Association Nutrition Committee," *Circulation*, vol. 114, pp. 82-96, 2006.
- [52] W. Mertz, "A balanced approach to nutrition for health: the need to biologicallyessential minerals and vitamins," *J. Am. Diet. Ass.*, vol. 94, pp. 1259-1262, 1994.
- [53] Report of the Director-General, "The World Health Report", World Health Organization, Geneva, Switzerland, 1997.
- [54] H.B. Hubert, M. Feinleib, P.M. McNamara and W.P. Castelli, "Obesity as an independent risk factor for cardiovascular disease: a 26- year follow-up of participants in the Framingham Heart Study," *Circulation*, vol. 67, pp. 968-997, 1983.
- [55] Centers for Disease Control and Prevention, "Prevalence of smoking among adults – United States, 2004," *Morbidity and Mortality Weekly Report*, vol. 54, No. 44, pp. 1121-1124, 2005.
- [56] *Composition of Foods Raw, Processed, Prepared, USDA National Nutrient Database for Standard Reference, Release 19*, U.S. Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Nutrient Data Laboratory, Beltsville, MD, August 2006.
- [57] G. Larsen and J. Conallen, "Engineering web-based systems with UML assets," *Annals of Software Engineering*, vol. 13, pp.203-230, 2002.
- [58] J. Conallen, *Building Web Applications with UML*, Reading MA: Addison Wesley, 2000.

- [59] D. Geary and C. Horstman, *Core Java Server Faces*, Upper Saddle River: Prentice Hall PTR, Sun Microsystems Press, 2004.
- [60] McGraw Hill/Osborn, "Introduction to JavaServer Faces, part 1," March 8, 2004, <http://www.devshed.com/c/a/Java/Introduction-to-JavaServer-Faces-1>
- [61] McGraw Hill/Osborn, "Introduction to JavaServer Faces, part 2," March 15, 2004, <http://www.devshed.com/c/a/Java/Introduction-to-JavaServer-Faces-Part-2/>
- [62] H. Bergsten, "Improving JSF by dumping JSP," June 9, 2004, <http://www.onjava.com/pub/a/onjava/2004/06/09/jsf.html>
- [63] R. Barcia, "JavaServer Faces (JSF) versus Struts," October 19, 2004, <http://websphere.sys-con.com/read/46516.htm>
- [64] Oracle JDeveloper Documentation, "About the lifecycle of a JSF page," November 15, 2007, http://www.oracle.com/webapps/online-help/jdeveloper/10.1.3/state/content/navId.4/navSetId._/vtTopicFile.jsf_apps%7Csf_create%7Csf_ajsfpagelifecycle%7Ehtml/
- [65] G. Anderson and P. Anderson, *Java Studio Creator Field Guide*, Upper Saddle River, NJ: Prentice Hall PTR, Sun Microsystems Press, 2004.
- [66] J. Brittain and I. Darwin, *Tomcat: the Definitive Guide*, Sebastopol: O'Reilley Media, 2003.
- [67] P. Dubois, *MySQL Cookbook*, Sebastopol: O'Reilley Media, 2002.
- [68] E. Bonder, "Basic Calendar," October 30, 2000, <http://javaboutique.internet.com/BasicCalendar/>
- [69] USDA Economic Research Service, "Major trends in U.S. food supply 1909-99," *Food Review*, vol. 24, No. 1, pp. 8-16, 2000.
- [70] USDA Economic Research Service, "Nutrient Content of the U.S. Food Supply: Nutrients and other food components, 1909 to 2004," Data Set, February 2007.
- [71] M. Cloutier and E Adamson, *The Mediterranean Diet*, New York: Avons Books, 2004.
- [72] K.K. Saunders, *The Vegan Diet as Chronic Disease Prevention: Evidence Supporting the New Four Food Groups*, Hemdon, VA: Lantern Books, 2003.

- [73] V. Messina, R. Mangels, and M. Messina, *The Dietitian's Guide to Vegetarian Diets*, Sudbury, MA: Jones and Bartlett Publishers, 2004.
- [74] J. Stepaniak, *Being Vegan*, Lincolnwood, IL: Lowell House/Contemporary Publishing Group, 2000.
- [75] W. Herrmann, H. Schorr, R. Obeid, and J. Geisel. "Vitamin B-12 status, particularly holotranscobalamin II and methylmalonic acid concentrations, and hyperhomocysteinemia in vegetarians," *American Journal of Clinical Nutrition*, vol. 78, pp. 131-136, 2003.
- [76] A.C. Antony, "Vegetarianism and vitamin B-12 (cobalamin) deficiency," *American Journal of Clinical Nutrition*, vol. 78, pp. 3-6, 2003.
- [77] Centers for Disease Control and Prevention, "Prevalence of heart disease - United States, 2005," *Morbidity and Mortality Weekly Report*, vol. 56, No. 06, pp. 113-11, 2007.
- [78] *Health, United States, 2002 with Chartbook on Trends in the Health of Americans*, National Center for Health Statistics, Hyattsville, MD, DHHS Publication No 1232, 2002.