Uppsala University

Laboratory Information Management Systems

Project

" Combat Colorectal Cancer "
" *CCC* "

Efthymia Chantzi

Email: Efthymia.Chantzi.0787@student.uu.se

April 3, 2015

Abstract

This project deals with the development of a web application that pertains to a simple and accurate risk prediction tool, with reference to colorectal cancer. Colorectal cancer comprises the third most common type of cancer and the fourth most common cause of deaths worldwide. However, there are high chances of treatment and survival in cases of early diagnosis. For this reason, "Gastro" clinic in cooperation with its specialists decided to expand their website by incorporating the "CCC"[1] online diagnosis tool. Users from all over the world are able to visit the platform, register in the system and perform the colorectal cancer risk test, which is provided as a simple and user-friendly questionnaire. In this way, they become familiar with their risk of developing colorectal cancer. Moreover, there is the opportunity to contact one of the clinic's doctors and ask for further diagnosis, consultation and advice. Previous and current activity is stored and accessed through personal accounts, so that individuals can keep track of their risk and health status. Statistical representations concerning the prevalence of the disease, with respect to fundamental factors, such as age, gender and nationality, are available not only for members and specialists of "Gastro" clinic, but also for researchers and visitors around the globe, who are interested in capturing and analyzing referenced information. The simulation of a recently approved prediction model together with its validation on predefined pathological cases, showed that the prediction accuracy is reliable and high. Consequently, the above features render this web application suitable in raising public awareness, investigating the risk of suffering from colorectal cancer and thus, facilitating early diagnosis.

---

[1]Combat Colorectal Cancer

Contents

List of Figures

4

List of Tables

1

Introduction

It is important to outline the bedrocks of this work, in order to render readers familiar with the basic concepts. However, followers who feel confident and have solid background on the anatomy of human gastrointestinal tract, may sidestep this information and continue with the following chapters.

1 Human Gastrointestinal Tract

The *Human Gastrointestinal Tract*, also known either as *GI tract*, or *GIT*, is a network of organs responsible for consuming and digesting food, absorbing nutrients, as well as expelling waste. The GI tract consists of the stomach and intestines. It is divided into two parts; *upper* and *lower*.
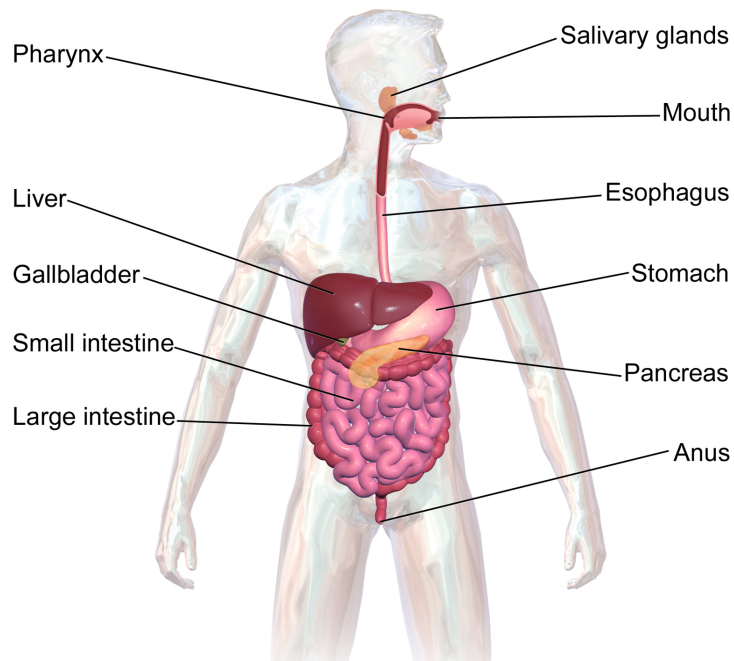
Figure 1.1: Human Gastrointestinal Tract - GIT [1]

2 Colorectal Cancer

Colon cancer is cancer of the large intestine[1], the lower part of humans' digestive system. *Rectal* cancer is cancer of the last several inches of the colon. Together, they are often referred to as **colorectal** cancer, also known as "CRC".

2.1 Pathology

Most cases of CRC begin as small, benign clumps of cells, called adenomatous polyps. Over time, some of these polyps become malignant tumors. Hence, the earlier the cancer is diagnosed, the better the chances for treatment and survival. Many people experience no symptoms in the early stages of the disease. However, when symptoms appear, they are likely to vary. The most prevalent include weakness, fatigue, unexplained weight loss and blood in the stool. CRC comprises the third most common form of cancer and the fourth most usual cause of deaths provoked by cancers, worldwide.
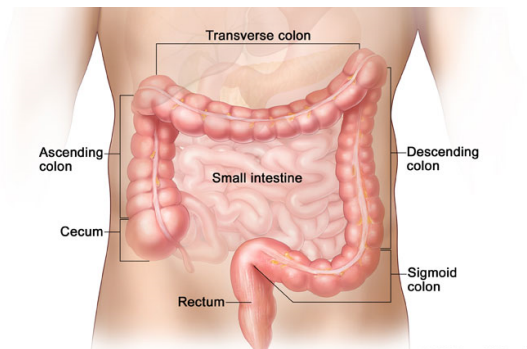
Figure 1.2: Lower GIT [2]

2.2 Risk Factors

It is important to mention that the vast majority of colorectal cancers are provoked by specific lifestyle factors along with increasing age. The main factors that are strictly associated with high risk of developing CRC include diet, obesity, smoking, alcohol consumption, lack of physical activity, family history and other diseases, such as diabetes or inflammatory bowel disease. A little improvement in the lifestyle is adequate to decrease significantly the risk of suffering from colorectal cancer.

2.3 Worldwide Precautions

Huge amounts of money are spent on CRC treatment each year, in all the prominent and developed countries, around the globe. Not only does early detection decrease the death rate, but it also reduces the expenses related to the cure of this cancer at higher stages, including chemotherapy and X-ray radiation.

Therefore, there is a strong need for social awareness regarding the necessity of healthier lifestyles, especially among populations that are considered to be at high risk. Healthy habits, such as balanced diet, intense physical activity and regular CRC screening tests should be promoted. In this way, social awareness is raised and medical prevention is enhanced.

---

[1]colon

3   Initiatives

"Laboratory Information Management Systems" [2] course was a great opportunity to acquaint the required technologies, in order to develop a software-based management system, which supports and facilitates the operations of a modern laboratory environment. Provided the high occurrence and severity of colorectal cancer, an online prediction tool enriched with the opportunity for immediate medical contact would be a significant contribution in the field of early diagnosis and prevention of "CRC". The ideal sector for hosting such a prediction tool, would be the website of a well known gastroenteric clinic with experience and expertise in colorectal cancer cases. Consequently, this project is implemented as a clinic-based system, so as to ensure reliability, security, confidentiality and support from specialists. In this way, the system could be further extended in a large-scale interactive platform with doctors and researchers from all over the world.

---

[2]LIMS

Description & Requirements

"***Combat Colorectal Cancer***"[1] constitutes a clinical center-based online diagnosis tool. More precisely, it is a web platform incorporated into the website of a gastroenteric clinic. Users are able to access the online "CCC" platform, perform the corresponding risk test and be aware of their risk, concerning the development of colorectal cancer. Further assessment of the result from the clinic's experts provides the opportunity for immediate consultation, medical advice and support. Chances are that people would find more convenient to perform a simple but reliable self-diagnosis test offered by the website of a clinic rather than visiting a physician, especially when there are not any apparent symptoms or signs of the disease.

1 Business Case

"Gastro" is a gastroenteric clinic, which is situated in the city of Uppsala, in Sweden. It is dedicated to deal with cancers of GI Tract, since 1995. More than 500 physicians, scientists and researchers share their expertise, so as to empower people confront with health problems, related to malfunctions of their digestive system. As most well-established clinics nowadays, "Gastro" maintains its own website. However, its owners have decided to expand their existent web platform by incorporating an innovative and accurate colorectal cancer prediction test. They feel that people must be able to figure out instantly using their laptops, tablets or smartphones, if they belong to high risk groups and take precautions. Their request depends on the fact that people should be self-motivated to gain knowledge, discover their risk and improve their lifestyle in terms of prevention and elimination of the disease. A carefully designed and charming user-interface that fulfills all the necessary requirements, can give a head start towards this direction.

Figure 2.1*: "Gastro clinic" in Uppsala*

---

[1]abbreviation: "CCC"

2 Developers and Development Method

Efthymia is the responsible developer for the implementation of the customer's request; "CRC"[2] online prediction tool. The customer; "Gastro" clinic, requires the completion of the project within two months; from the beginning of February to the end of March 2015.

In terms of "CCC", the most important functionality is the implementation of the online prediction test, which must be performed successfully by the users, who visit and get registered to the "Gastro" clinic's website. If this functionality is sufficiently provided until the middle term presentation, the application is going to be expanded further. Records of the system along with detailed timeplan are included in the following chapters, where each part is presented thoroughly.

3 Requirements

"CCC" must be developed according to the customer's requirements. Only then, the system can be approved. It is true that the list of the requirements is quite long, but the most fundamental are summarized in table 2.1.

Property	Description
Security	Users' information, test inputs and results must be stored safely only in personal accounts. This kind of information must not be accessed by anyone else, apart from users who own a particular personal account.
Reliability	Test results must be estimated reliably and validated. When it comes to risk of suffering from colorectal cancer, the result must be accurate, consistent and verified. It is very important to recognize the severity of this prediction test and adopt a responsible behavior.
Scalability	The system must be able to perform well with increasing number of registered users. As more users become members of the web platform and make use of it, the system must be able to accommodate this growth. It must be also customizable and extensible.
Outstanding user-interface	The user-interface must be attractive, simple to use and self-explained, at the same time. There exist different target groups of varying age, education and nationality. Therefore, the user-interface must be adjusted to serve all these different goals.

Table 2.1: Customers' general requirements for "CCC"

---

[2]Colorectal Cancer

3

System Architecture & Implementation

"CCC", like any other proper LIMS software, constists of three different main layers:

(a) Database design

(b) Web server

(c) User-interface

This chapter elaborates on each one of the aforementioned layers and sheds light on how they are designed and implemented.

1   Users

First and foremost, it is very important to state the different categories of users, for whom this LIMS application has been developed. There exist **three** different categories of users for "CCC". More precisely,

(i) **public users** $\longrightarrow$ people, who get registered in the system through "CCC" web interface, and maintain their own personal account. After successful registration, they are able to access the system via login, perform the prediction test, interact with specialists and keep records of their health status throughout time.

(ii) **doctors** $\longrightarrow$ specialists, who are active members of "Gastro" clinic. Apart from their activity in the clinic, they have been provided credentials to access the "CCC" system and thus, handle requests for medical consultations from public users. They also have personal accounts with options of performing the risk prediction test if they wish to, and keeping records of their health progress with respect to time and improved lifestyle.

(iii) **visitors** $\longrightarrow$ individuals of unspecified nature; they are, in general, visitors of the website, who aim at gaining knowledge about colorectal cancer, services of "Gastro" clinic, rate and occurrence of the disease based on the clinic's statistical measurements. Consequently, they may be people, who wish to become familiar with the concept of CRC, or researchers that study the expansion of this cancer around the world. In this case of pure information, users are just visitors and they do not need to access the system and create personal accounts.

2   Use Cases

In the previous section (1), the three different types of users were presented. Use cases describe in detail the allowed activities and how they are efficiently handled by the system. Elaborated use case diagrams as wells as detailed descriptions of them, are following.

2.1 Diagrams

2.1.1 Public Users
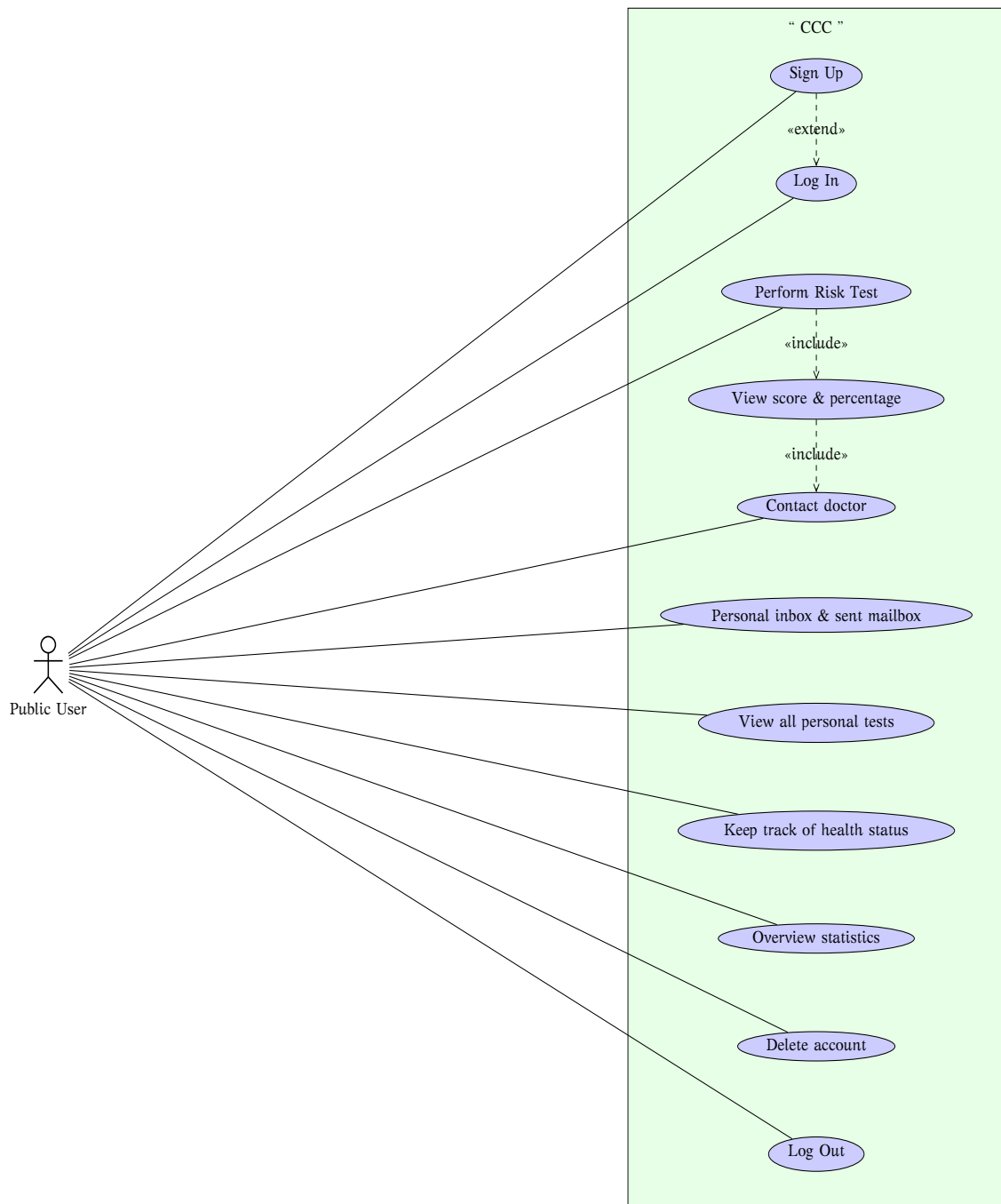
Figure 3.1: Use cases ⟶ Public User

2.1.2 Doctors
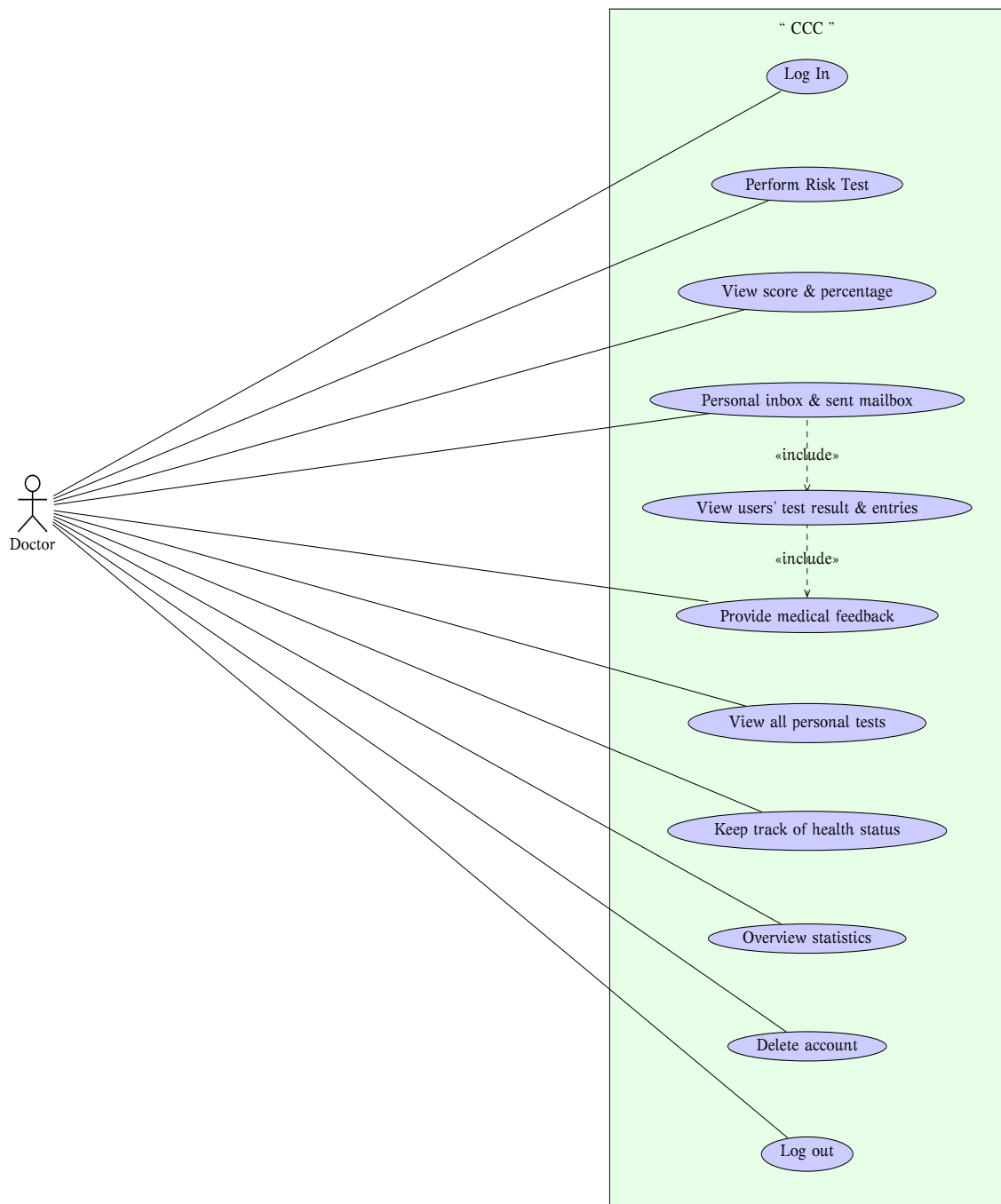
Figure 3.2: Use cases ⟶ "Gastro" Clinic's Doctor

2.1.3 Visitors
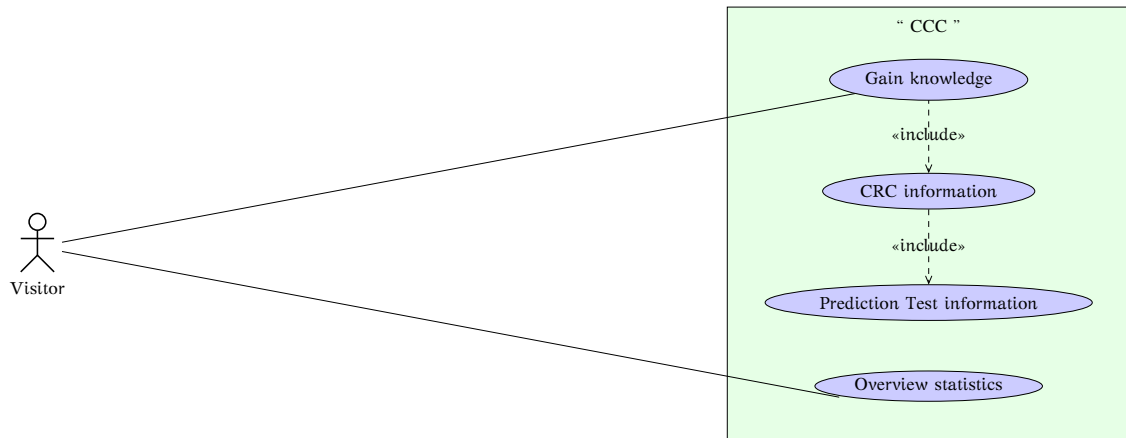
Figure 3.3: *Use cases* ⟶ *"CCC" website's visitor*

2.2 Descriptions

2.2.1 Sign Up

Use Case	Sign Up
Date Created:	February 10, 2015
Last Update:	March 2, 2015
Primary Actor:	• public user
Description:	This use case pertains to public users, who wish to become registered members of "CCC", obtain credentials and maintain their own personal account.
Trigger:	Clickable button "Sign Up" on the main menu of "CCC".
Pre-conditions:	Access to the main page of "CCC" web platform
Assumptions:	Each user owns a unique pair of credentials; *username, password*
Post-conditions:	• welcome message for successful registration • log in

Normal flow:

1. User clicks on the button "Sign Up" from the navigation bar of the "CCC" home page.
2. Direction to the corresponding page containing the sign up form.
3. User fills in the given fields; username, password, retype password, firstname, lastname, gender, email, country. All of them are mandatory except for email.

4. User clicks on the "Sign Up" button at the end of the form.

Success Scenario	• If the form is completed correctly and no errors exist, the user is inserted successfully into the DBMS and a welcome with the newly created username is displayed.

Failure Scenario:

1. One or more of the required fields are left blank:
    (i) system shows failure message
    (ii) user is prompted to re-submit the form by filling in all the required fields, which are highlighted
2. Username and/or password already exist in the DBMS:
    (i) system shows failure message
    (ii) user is prompted to re-submit the form by entering an other username and/or password
3. Weak password; less than 5 characters long
    (i) system shows failure message
    (ii) user is prompted to re-submit the form by entering a new longer password
4. Mismatched passwords; "password", "retype password" fields
    (i) system shows failure message
    (ii) user is prompted to re-submit the form carefully, so that the two fields match
5. Invalid data; firstname/lastname consists of numbers, email not in the conventional format
    (i) system shows failure message
    (ii) user is prompted to re-submit the form carefully, by entering only letters in firstname/lastname field, and typing the email in the right format; something@domain.suffix

2.2.2 Log In

Use Case	Log In
Date Created:	February 10, 2015
Last Update:	February 12, 2015
Primary Actor:	- public user - doctor
Description:	This use case pertains to public users or clinic's doctors, who wish to access their personal account by providing their credentials; *username, password.*
Trigger:	Clickable button "Log in" on the main menu of "CCC".
Pre-conditions:	Access to the main page of "CCC" web platform.
Assumptions:	Each user owns a unique pair of credentials; *username, password.*
Post-conditions:	Access to the personal page of each user. There is a navigation bar which guides the user to execute the actions of his/her preference.

Normal flow:

1. User clicks on the button "Log In" from the navigation bar of the "CCC" home page.
2. Direction to the corresponding page containing the log in form.
3. User fills in the given fields; username and password. Both are required.
4. User clicks on the "Log In' button at the end of the form.

Success Scenario	If the correct combination of credentials is provided, log in of the user is successfully performed.

Failure Scenario:

1. One or more of the two required fields; username, password, are left blank:
   - (i) system shows failure message
   - (ii) user is prompted to re-submit the form by filling in both required fields, which are highlighted
2. Invalid username and/or password; meaning that they are not stored in the corresponding table of DBMS:
   - (i) system shows failure message
   - (ii) user is prompted to re-submit the form, by entering the correct credentials

Frequency of Occurrence:	Each time user(s)/doctor(s) wishes to log in to their personal account.

2.2.3 Perform Colorectal Cancer Risk Prediction Test

Use Case	CRC Risk Prediction Test
Date Created:	February 13, 2015
Last Update:	February 20, 2015
Primary Actor:	• public user • doctor
Description:	This use case pertains to public users or clinic's doctors, who wish to take the CRC risk prediction test and be aware of their risk, concerning the development of colorectal cancer.
Trigger:	Clickable button "Test" on the navigation bar of their personal page.
Pre-conditions:	Successful log in to their personal account.
Assumptions:	NA
Post-conditions:	• display of risk score and percentage • prompt for further contact with one of the clinic's specialists

Normal flow:

1. User performs successful log in to his/her personal accounted, as indicated by the aforementioned use case.
2. User clicks the button "Test" from the navigation bar of his/her personal page.
3. Transition to the page that contains the risk test form with all the required entries.
4. User fills in the questionnaire with respect to the instructions given.
5. User clicks on the "Proceed" button at the end of the questionnaire.

Success Scenario	• If all required fields are completed, the risk score along with the respective percentage are displayed. • Prompt for further interaction with a doctor from "Gastro" clinic.

Failure Scenario:

1. One or more entries of the risk questionnaire are left blank:
   (i) system shows failure message
   (ii) user is prompted to re-submit the questionnaire by filling in all fields, which are highlighted
2. Invalid input; numbers beyond the allowed limits, insertion of characters when only numbers are required.
   (i) system shows failure message
   (ii) user is prompted to re-submit the form, by entering the correct type of entries

Frequency of Occurrence:	Each time user(s)/doctor(s) wishes to take the CRC prediction test.

2.2.4 Contact Doctor

Use Case	CRC Risk Prediction Test
Date Created:	February 15, 2015
Last Update:	February 15, 2015
Primary Actor:	• public user
Description:	This use case pertains to public users, who wish to ask for further consultation after having performed the risk test. The user selects the doctor of his/her own preference and he/she has the opportunity to send a personal text message apart from the test inputs and result. This message is forwarded to the respective doctor's account as an inbox message.
Trigger:	Clickable button "Diagnosis" on the bottom of CRC's risk score page.
Pre-conditions:	Successful submission of the risk questionnaire.
Assumptions:	NA
Post-conditions:	• display of contact form with a textarea field. However, this is optional. It can be used in case the user wants to type a personal message to doctor. • list of available doctors. The user must choose exactly one of them, so as to send his/her request for medical advice.

Normal flow:

1. User clicks the button "Diagnosis" at the bottom of the test's result page.
2. Transition to the corresponding contact form page. There is the list of doctors to choose from and an optional textarea, if the users want to include a personal written message.
3. User clicks on the "Send" button at the end of the contact form, when he/she has chosen a doctor from the check list and/or has written the message.
4. Verification that the message has been successfully sent to the respective doctor.
5. The sent message can be retrieved through the personal sent mailbox from the navigation bar.

Success Scenario:	User has chosen a doctor from the clinic's checklist.

Failure Scenario:

1. None of the doctors is selected but the button "Diagnosis" is pressed:
   (i) system shows failure message
   (ii) user is prompted to choose a doctor from the available list and re-send the request

Frequency of Occurrence:	Each time a registered user wishes to take the CRC prediction risk test and decides that he/she should consult a doctor from the clinic.

2.2.5 Personal inbox & sent mailbox

Use Case	**Handle personal inbox & sent mailbox**
Date Created:	March 2, 2015
Last Update:	March 10, 2015
Primary Actor:	• public user • doctor
Description:	This use case refers to the interaction between public users and doctors via written messages. It arises from the fact that registered public users are able to contact one of the clinic's doctors for further consultation, each time after having performed the CRC risk test. The submitted risk test, including the risk factors and score, along with an optional written message is sent to the selected doctor's account, as an inbox message. Similarly, each registered user receives inbox messages from doctors, when they replay back and provide feedback. Once inbox messages are read by both sides, they are transferred to the sent mailbox of each account, where they can be retrieved and seen at any time.
Trigger:	Clickable buttons "Inbox", "Sent" from the drop down list "Messages" in the navigation bar of personal account.
Pre-conditions:	Successful log in to personal account.
Assumptions:	NA
Post-conditions:	• list of unread/inbox messages as an interactive HTML table. Each row corresponds to a different message and by clicking on it, the user is able to retrieve and read it. • list of sent messages as an interactive HTML table. Each row corresponds to a different message and by clicking on it, the user is able to retrieve and read it.

Normal flow:

1. User clicks on the drop down list "Messages" in the navigation bar of personal page. Then, a specific category is selected by either clicking "Inbox" or "Sent" messages.
2. Transition to the corresponding maibox; *inbox* or *sent*. Messages can be retrieved separately by clicking on each row of the mailbox table.

Success Scenario:	Messages, either inbox or sent are retrieved and read. When a message is sent successfully after the button "Send" has been clicked, a confirmation message is displayed.

Failure Scenario:	NA. In case there are no available inbox or sent messages, the respective mailbox table informs the user/doctor that it is empty.
Frequency of Occurrence:	It depends on the number of inbox messages and how many times the user/doctor wishes to overview his/her sent mailbox.

2.2.6 Medical Feedback

Use Case	Provide medical feedback
Date Created:	March 5, 2015
Last Update:	March 7, 2015
Primary Actor:	• doctor
Description:	This use case pertains to doctors, who receive consultation requests from public users. In turn, they have to inspect their test inputs and score, and reply back as soon as possible.
Trigger:	Clickable row in the interactive inbox table.
Pre-conditions:	Successful log in to personal account.
Assumptions:	Doctor has new/unread messages from public users, who ask for further medical advice.
Post-conditions:	Medical response is sent back to public users as an inbox messages.

Normal flow:

1. Doctor clicks on the drop down list "Messages" in the navigation bar of personal page and selects the option "Inbox (*# unread messages*)".
2. Transition to the personal interactive inbox repository. Unread messages are included in separate rows from the most to the least current.
3. Doctor clicks on the row containing the message he/she wants to read.
4. Doctor views the message along with the risk factors related to the test that the user has performed.
5. Doctor writes his/her own response in the appropriate textarea at the bottom of the form.
6. Doctor clicks on the button "Send" in the right side of the response textarea.
7. The sent message can be retrieved by clicking on the field "Sent" in the navigation bar of personal page.

Success Scenario:	A confirmation message is displayed when the message has been sent successfully. It can also be retrieved through personal sent repository.

Failure Scenario:

1. The button "Send" is pressed but the doctor has omitted to write his/her own message:
   (i) system shows failure message
   (ii) doctor is prompted to include his/her feedback in the appropriate text area, which is highlighted

Frequency of Occurrence:	It depends on the number of inbox messages that the doctor has received.

2.2.7 Personal risk tests repository

Use Case	**View all personal risk tests**
Date Created:	March 10, 2015
Last Update:	March 10, 2015
Primary Actor:	• public user • doctor
Description:	This use case refers to public users and doctors, who wish to have a look at the history of the CRC risk tests that they have performed so far.
Trigger:	Clickable field "Previous Tests" from the drop down list "Test" in the navigation bar of personal account.
Pre-conditions:	Successful log in to personal account and having performed at least one CRC risk test so far.
Assumptions:	NA
Post-conditions:	list of performed tests as an interactive HTML table. Each row corresponds to a different test including date and time as well as the percentage of the risk score.

Normal flow:

1. User clicks on the drop down list "Test" in the navigation bar of personal page and selects the field "Previous Tests".
2. Transition to the corresponding CRC risk test repository.
3. Risk factors can be retrieved separately by clicking on each row that responds to a specific CRC risk test.

Success Scenario:	Previous tests; inputs, risk factors and risk score are retrieved and available.
Failure Scenario:	NA. In case there are no performed CRC risk tests, the respective repository informs the user/doctor that it is empty by displaying the appropriate message.
Frequency of Occurrence:	It depends on how many times the user/doctor wishes to overview his/her performed CRC risk test(s).

2.2.8 Health Status

Use Case	Keep track of health status
Date Created:	March 10, 2015
Last Update:	March 10, 2015
Primary Actor:	• public user • doctor
Description:	This use case refers to public users and doctors, who wish to keep track of their health status concerning CRC risk. It is associated with a bar chart that depicts the risk score % with respect to time, so that users can check their progress over time.
Trigger:	Clickable field "Progress" from the drop down list "Test" in the navigation bar of personal account.
Pre-conditions:	Successful log in to personal account and having performed at least one CRC risk test so far.
Assumptions:	NA
Post-conditions:	Bar chart, where the horizontal axis contains the date of each performed risk test, while the vertical depicts the corresponding risk score %.

Normal flow:

1. User clicks on the drop down list "Test" in the navigation bar of personal page and selects the field "Progress".
2. Illustration of the corresponding bar chart. By clicking on each bar, information about the date and score is displayed.

Success Scenario:	Visualization of bar chart.

Failure Scenario:

1. The button "Progress" is clicked but the user has not performed any CRC risk test yet.
   (i) system shows information message
   (ii) user is prompted to perform at least one test in order to be able to visualize the score as a bar and at least two, so as to view his/her progress

Frequency of Occurrence:	It depends on how many times the user/doctor wishes to overview his/her health status.

2.2.9 Overview statistics

Use Case	Overview statistics
Date Created:	March 12, 2015
Last Update:	March 16, 2015
Primary Actor:	• public user • doctor • visitor
Description:	This use case refers to public users doctors and visitors, who wish to be informed about the overview statistics of "Gastro" clinic regarding colorectal cancer.
Trigger:	Clickable fields "Gender", "Country" from the drop down list "Statistics" in the navigation bar of personal account.
Pre-conditions:	NA, since overview statistics are available either from the home page or through personal accounts. They are made for research purposes and this is why they can be accessed by all three kinds of users.
Assumptions:	NA
Post-conditions:	• Scatter chart, where the CRC risk score % is depicted with respect to age among male and female population of "CCC" DBMS. Individuals are illustrated as interactive bubbles, whose size is relative to the risk score. Different color among males and females render the discrimination possible. As the populations grows, the distribution becomes more apparent and reliable. • Worldwide map, where all countries, related to "CCC" users that have performed the risk test, are illustrated, with the help of a pseudocolor scale. In each country the mean value of all the risk scores associated with their people, is assigned.

Normal flow:

1. User clicks on the drop down list "Statistics" in the navigation bar of either home or personal page and selects the fields "Gender", "Country".
2. Illustration of scatter chart. By clicking on each bubble, information about age, gender and risk score. Each bubble represents an unrecognized individual that has performed the "CCC" risk prediction test.
3. Illustration of map chart. By clicking on each country, its name, risk score % and position in the pseudocolor scale are displayed.

Success Scenario:	Visualization of bar and map charts.
Failure Scenario:	NA

Frequency of Occurrence: It depends on how many times the users wishes to study the overview statistics estimated by "CCC" DBMS.

2.2.10 Delete account

Use Case	Delete account
Date Created:	March 14, 2015
Last Update:	March 14, 2015
Primary Actor:	• public user • doctor
Description:	This use case refers to public users and doctors, who wish to delete their account from the "CCC" system.
Trigger:	Clickable field "Delete Account" from the drop down list "Profile *username*" in the navigation bar of personal account.
Pre-conditions:	Successful log in to personal account.
Assumptions:	NA
Post-conditions:	transition to "CCC" home page as a visitor, with limited privileges that correspond to this category.

Normal flow:

1. User clicks on the drop down list "Profile *username*" in the navigation bar of personal page and selects the field "Delete Account".
2. The user clicks on the button "Confirm" if he/she wants to verify the decision and wants to proceed with the deletion of the personal account.
3. Redirection to the home page as a simple visitor of the website.
4. The user clicks on the button "Cancel", in case he/she has changed his/her mind and wants the maintain the personal account.
5. Redirection back to personal profile page.

Success Scenario:	Deletion of the personal account.
Failure Scenario:	NA
Frequency of Occurrence:	Only once.

2.2.11 Log out

Use Case	Log out
Date Created:	March 14, 2015
Last Update:	March 14, 2015
Primary Actor:	• public user • doctor
Description:	This use case refers to public users and doctors, who wish to log out from their personal account and terminate their current session.
Trigger:	Clickable field "Log Out" from the drop down list "Profile *username*" in the navigation bar of personal account.
Pre-conditions:	Successful log in to personal account.
Assumptions:	NA
Post-conditions:	transition to "CCC" home page as a visitor, with limited privileges that correspond to this category.

Normal flow:

1. User clicks on the drop down list "Profile *username*" in the navigation bar of personal page and selects the field "Log Out".
2. Redirection to the home page as a simple visitor of the website.

Success Scenario:	Termination of current session.
Failure Scenario:	NA
Frequency of Occurrence:	Unlimited number of times, as long as a personal account exists.

3   Database Design

The developed database for this LIMS web application consists of:

  (i) six tables

 (ii) two restricted views; two out of six tables

(iii) one stored procedure, which estimates the colorectal cancer risk score based on a validated prediction model (4).
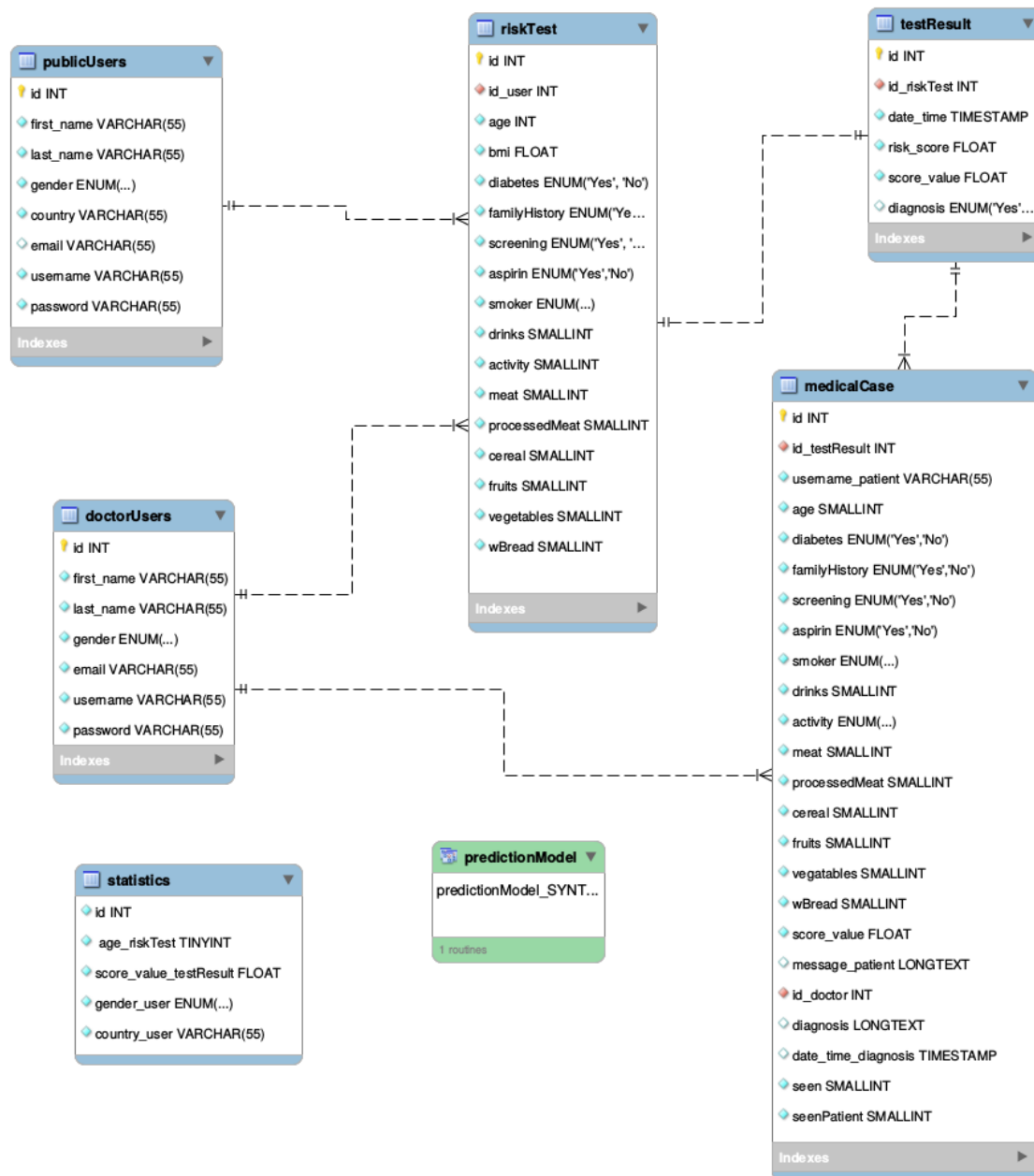
3.1   Schema

Figure 3.4: "CCC" database schema

3.2 Functionality

The functionality in the "CCC" DBMS is guaranteed by a reasonable and consecutive flow of information among the different tables.

First and foremost, there are two different categories of registered users; *public users* and *doctor users*. Looking carefully at figure 3.4, it is possible to locate the two tables related to them; *publicUsers* and *doctorUsers*. Secondly, the table *riskTest* is strictly associated with the risk factors of the test that the user is required to fill in the specially designed questionnaire. Users are able to perform a great number of risk tests, but at relatively long period between them, since even a little improvement in the lifestyle needs time to yield good results.

Furthermore, *testResult* is the table concerning the risk score as well as the risk score percentage, which is being estimated by a simulated prediction model triggered by the respective stored procedure; *predictionModel*. One risk test corresponds to one result. The prediction model is going to be explained explicitly in the next section; section 4.

As already mentioned in 2.2.4, there exists the use case "Contact Doctor". Registered users, who have taken the test, they have the opportunity to contact one of the clinic's specialists and ask for their advice and help. The table *medicalCase* is implemented to serve this purpose. Actually, it is a restricted view over tables *publicUsers*, *riskTest*, *testResult*, and *doctorUsers*. Briefly, it contains the message as well as the risk factors of the user's test that is sent to a specific doctor. In this way, the responsible doctor can gain access only to test data and message, without being able to recognize the individual. The provided diagnosis is also stored in this table, and the user is able to access it.

Finally, there is the table *statistics*, which is used to keep the overview statistics that are associated with "CCC". Personal data of the users are confidential, which means that they must be accessible only by them through their registered personal accounts. Hence, the risk score (%) along with other meaningful parameters, such as age, gender and country are stored in the restricted view "statistics". This is done each time a risk test is completed and the respective risk score is displayed. Even if the user deletes his/her account, these data remain in the "statistics" table. However, it should be stated that nobody can be recognized from the combination of age, gender, country and risk score. Therefore, security is guaranteed.

3.3 Tables

3.3.1 Public/Doctor Users

publicUsers	
Attributes:	personal data of users
Primary key:	id
Use case:	sign up, log in

Table 3.12: public users

doctorUsers	
Attributes:	personal data of doctors
Primary key:	id
Use case:	log in

Table 3.13: doctor users

It should be mentioned that the "sign up" procedure of doctors is handled by the DBA. In other words, "CCC" does not provide a separate "sign up" form for doctors. Since, for the time being, it is a clinic-based system, the registration of doctors is not provided by the user-interface itself.

3.3.2 CRC Prediction Risk Test

riskTest	
Attributes:	risk factors for CRC test
Primary key:	id
Foreign key:	id_user
Parent table:	publicUsers, doctorUsers
Use case:	perform CRC risk test

Table 3.14: *risk test questionnaire*
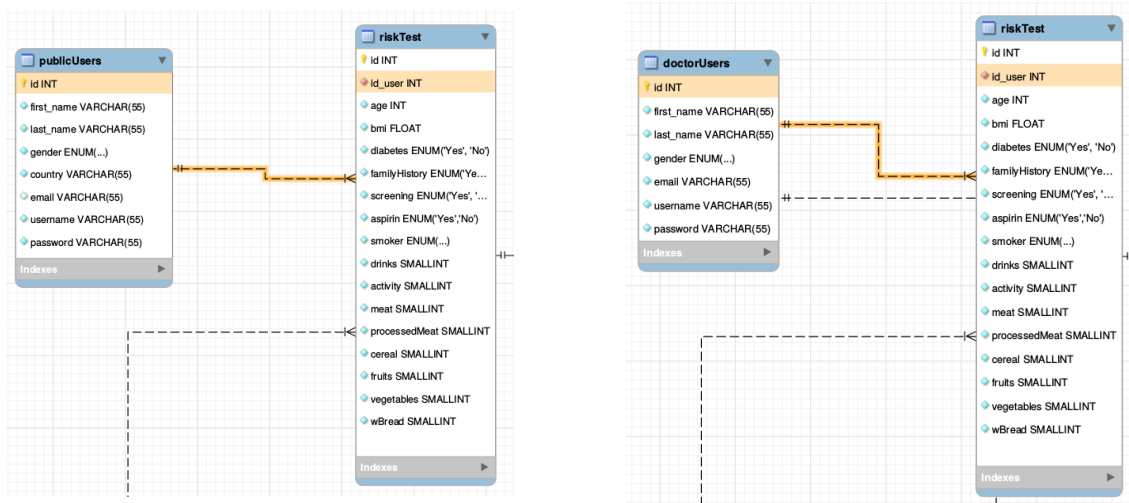
Figure 3.5: *relation view "publicUsers" - "riskTest", "doctorUsers" - "riskTest"*

3.3.3 Test Result

testResult	
Attributes:	risk factors
Primary key:	unique id
Foreign key:	id_riskTest
Parent table:	riskTest
Use case:	view & store risk score and percentage

Table 3.15: *risk test result*
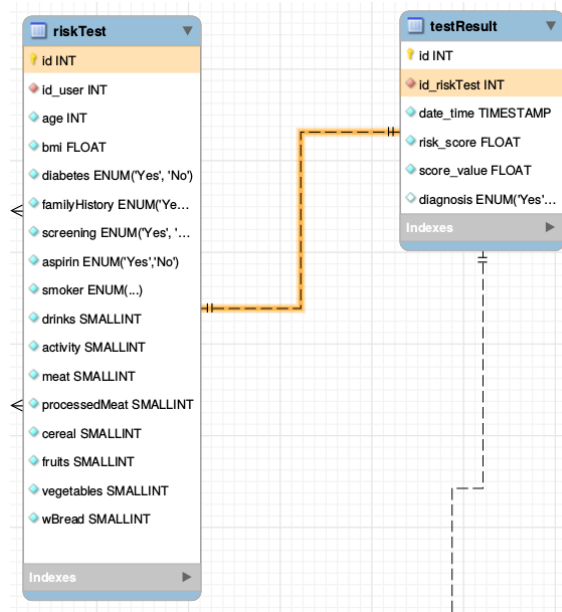
Figure 3.6: *relation view "riskTest" - "testResult"*

3.3.4 Medical Case

medicalCase	
Attributes:	risk factors, public user, doctor
Primary key:	id
Foreign key:	id_testResult, id_doctor
Parent table:	testResult, doctorUsers
Use case:	request/provide medical feedback

Table 3.16: *medical case*
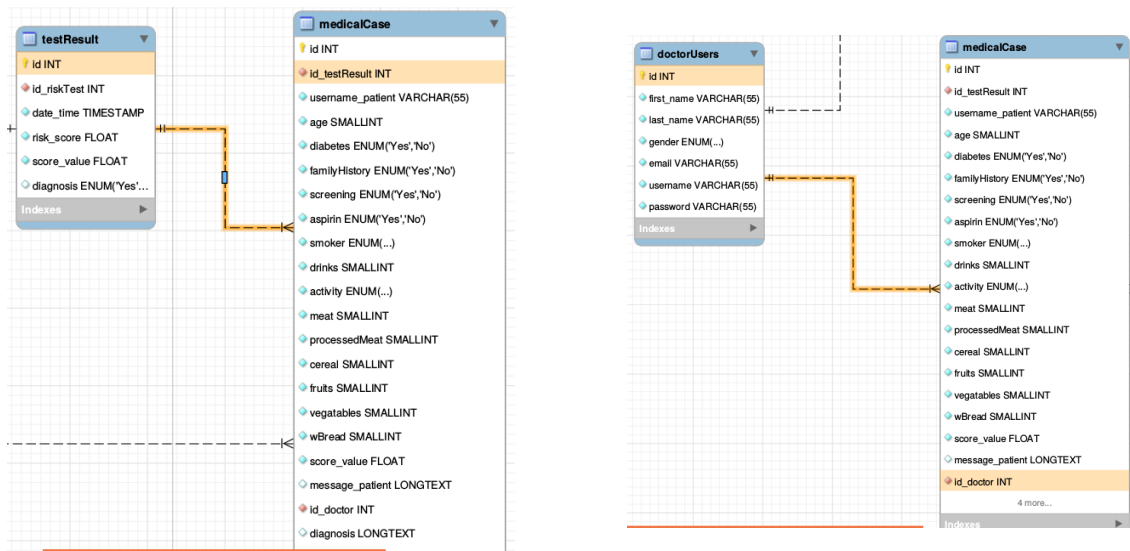
Figure 3.7: *relation view "medicalCase" - "testResult", "medicalCase - doctorUsers"*

3.3.5   Statistics

statistics	
Attributes:	score%, age, gender, country
Primary key:	id
Foreign key:	NA
Parent table:	NA
Use case:	overview statistics

Table 3.17: *overview statistics*

Figure 3.8: *statistics database table*

4 Prediction Model

The prediction model that is used for the estimation of risk score, concerning the development of colorectal cancer, is adopted by a recently published paper [3]. The underlying reason to simulate this prediction model is the fact that it is developed on real patients' data and has been validated. Moreover, it has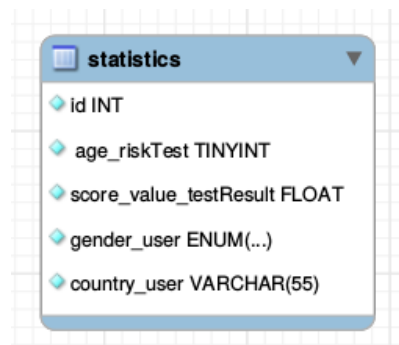 been published on August 2014, which indicates that it provides a recent and innovative perspective. It is important to mention that this paper provides two different models; *reduced* and *full*. In terms of "CCC", the latter one has been simulated, since it covers more risk factors and thus, it is more accurate than the reduced model.

4.1 Formulae

4.1.1 Pure Risk Score

The adopted prediction model constitutes a linear regression model. Consequently, the total risk score that responds to a submitted risk test is calculated by equation 3.1, which is implemented as a stored procedure in "CCC" DBMS.

As described by equation 3.1, there are 16 different coefficients; $(\beta_1, \cdots, \beta_{16})$, and 16 different risk factors either binary or continuous; $(\psi_1, \cdots, \psi_{16})$. Risk factors are included in the respective questionnaire of the user-interface, which is filled in by the user during the test procedure.

$$score(\beta, \psi) = \beta_1 \times \psi_1 + \cdots + \beta_{16} \times \psi_{16} \tag{3.1}$$

Table 3.18 illustrates the values and ranges of coefficients and risk factors, as well:

Full Model				
Risk factor (ψ_i)	Type	Coding	Coefficient (β_i)	Fitting
age	continuous	-	0.063	$age \times 0.063$
gender	binary	Female (1)	-0.283	$1 \times (-0.283)$
		Male (0)	-0.283	$0 \times (-0.283)$
BMI	continuous	-	0.018	$BMI \times 0.018$
diabetes	binary	Yes (1)	0.209	1×0.209
		No (0)	0.209	0×0.209
family history	binary	Yes (1)	0.076	1×0.076
		No (0)	0.076	0×0.076
CRC screening	binary	Yes (1)	-0.536	$1 \times (-0.536)$
		No (0)	-0.536	$0 \times (-0.536)$
aspirin use	binary	Yes (1)	-0.007	$1 \times (-0.007)$
		No (0)	-0.007	$0 \times (-0.007)$
smoker	categorical	Current (1)	0.254	1×0.254
		Former (1)	0.241	1×0.241
		No (1)	0	1×0
alcoholic drinks (#/day)	continuous	-	0.077	$drinks \times 0.077$
physical activity (hrs/day)	categorical	None (1)	0.043	1×0.043
		(1, 3.5] (1)	-0.098	$1 \times (-0.098)$
		(3.5 14] (1)	-0.249	$1 \times (-0.249)$
red meat (times/week)	binary	≥ 5 (1)	0.053	1×0.053
		< 5 (0)	0.053	0×0.053
processed meat (times/week)	binary	≥ 5 (1)	0.146	1×0.146
		< 5 (0)	0.146	0×0.146
cereal intake	binary	> 1 (1)	-0.088	$1 \times (-0.088)$
		None (0)	-0.088	$0 \times (-0.088)$
fruits (serves/day)	binary	< 2 (1)	0.077	1×0.077
		≥ 2 (0)	0.077	0×0.077
vegetables (serves/day)	binary	< 5 (1)	0.083	1×0.083
		≥ 5 (0)	0.083	0×0.083
wholemeal bread (serves/week)	binary	≥ 5 (1)	-0.007	$1 \times (-0.007)$
		< 5 (0)	-0.007	$0 \times (-0.007)$

Table 3.18*: Risk factors & co-efficients of the prediction model*

4.2 Risk Score Percentage

Apart from the pure risk score, which is calculated by the contribution of equation 3.1, the percentage that corresponds to this risk score, is also displayed and stored. It is usually more convenient and understandable to display the percentage rather than the pure risk score. The risk score percentage is gauged by the following formula:

$$\text{percentage}(\beta, \psi) = \left(\frac{\text{score}(\beta, \psi)}{\text{maximum}_{\text{score}}} \right) \times 100 \tag{3.2}$$

The value for the maximum obtained risk score is calculated by the sum of the *extreme* values for each one of the risk factors. By the term *extreme*, the values counting for the highest score for each risk factor are considered. Based on table 3.18 and top hish risk values, the maximum score is calculated as:

$$
\begin{aligned}
\textbf{maximum}_{\textbf{score}} = {}& \underbrace{100 \times 0.063}_{\text{100 years old}} + \underbrace{0 \times (-0.283)}_{\text{Male}} + \underbrace{50 \times 0.018}_{\text{50kg/m}^2} + \underbrace{1 \times 0.209}_{\text{Diabetes}} + \underbrace{1 \times 0.076}_{\text{CRC family history}} + \underbrace{0 \times (-0.007)}_{\text{No aspirin use}} \\
& + \underbrace{1 \times 0.254}_{\text{Current smoker}} + \underbrace{12 \times 0.077}_{\text{12 drinks/day}} + \underbrace{1 \times 0.043}_{\text{None activity}} + \underbrace{1 \times 0.053}_{\text{red meat} \geq 5} + \underbrace{1 \times 0.146}_{\text{processed} \geq 5} + \underbrace{0 \times (-0.088)}_{\text{no cereals}} \\
& + \underbrace{1 \times 0.077}_{\text{fruits} < 2} + \underbrace{1 \times 0.083}_{\text{vegies} < 5} + \underbrace{0 \times (-0.007)}_{\text{wbread} < 5} \\
& = \textbf{9.065}
\end{aligned}
$$

5   Development Platform

For the development of "CCC", **LAMP** has been used; an open source web stack platform, which uses **Linux** as the operating system, **Apache** as the web server, **MySQL** as the RDBMS[1] and **PHP** as the object-oriented scripting language.

5.1   Database Implementation Technology

For the implementation of "CCC" RDBMS, **MySQL 5.5** has been used. However, the database is also accessible through phpMyAdmin. It is very popular and convenient due to its web user interface. It is true that in cases of extreme workload, phpMyAdmin offers great help.

MySQL Workbench and especially **EER**[2] **Diagram Editor** has been used for the design of the database schema, which in this case is called *EER diagram*.

5.2   Languages

It is true that this project gives the opportunity to gain experience on several different languages that are used for web development. For the needs of "CCC" the used languages are summarized by the following list:

   (i) HTML5

   (ii) PHP

  (iii) CSS3

  (iv) JavaScript

   (v) JSON

5.3   Integrated Development Environment

NetBeans IDE 8.0.2 has been used in order to write the code in all the aforementioned five languages, required for the implementation of "CCC".

5.3.1   Bug Control

Debugging of php scripts has been performed using *XDebug debugger*, which works mainly with breakpoints. What is more, debugging of the JavaScript files has been offered by the *JavaScript debugger*. It should also be mentioned that unit tests on Javascript files were implemented using the *Jasmine* [4] testing framework. In this way, the increments were tested and validated before being incorporated and extended.

---

[1]Relational Database Management System
[2]Enhanced Entity Relationship

6 Technical Aspects

6.1 Error Handling

Error handling has been thoroughly presented in section 2.2, where each use case along with failure scenario and system's response are explained.

6.2 Security

Security is provided both on database and web server layer. First and foremost, as described in sections 3.1 and 3.2, the design of the database itself guarantees security, as far as the stored and retrieved data are concerned. Secondly, **MD5** cryptographic hash function has been used for the users' credentials, so as to guarantee data integrity.

In addition, **validation** and **sanitization** of all external inputs are performed in the middle-ware layer by using appropriate php filters. **SQL injection** attacks constitute a fundamental security problem, which is also taken into consideration. For these reasons, the following function has been developed:

```
function secure_data($data, $conn) {

        $data = trim($data);
        $data = strip_tags($data);
        $data = stripslashes($data);
        $data = htmlspecialchars($data);
        $data = mysqli_real_escape_string($conn, $data);

        return $data;

}
```

It should be stated, as well, that [POST] values are used instead of [GET], since the latter include user parameters in the URL, while the former put them in the body of HTTP request.

4

Results

The results of our LIMS application were presented in detail, during the final project presentation, on March 20, 2015. It is difficult to demonstrate the functionality of "CCC" by providing only screenshots of the developed user interface, but there is not any host server available. The screenshots of the "CCC" user interface can be found at:

https://drive.google.com/folderview?id=
0B4mMtieYoaNDfm96MFBHYWxCaO5hMVBlbExOZUt4bGRobm1Fd2NzZWVZVWF2ck9nelp5azQ&usp=
sharing

5

Conclusion & Future Work

The primary goal of the project has been achieved; "CCC" is functional. The website of "Gastro" clinic is equipped with an online colorectal cancer prediction tool, which helps people find out their impending risk of developing this disease, improve their lifestyle, keep health status records, interact with doctors and be informed about the rate of CRC occurence, worldwide. All these options are available, through a user-friendly and attractive interface, in conjunction with a secure and reliable database.

The available time for the implementation of the project was limited, which means that the LIMS application could not be a large-scale project by definition. However, there is a good basis for "CCC", which can be further expanded in later versions. It is true that there exist several aspects for improvement. A promising initiative for future work would be the conversion of "CCC" from a clinic-based center to an international platform, without borders, where doctors with expertise in colorectal cancer could cooperate and contribute to its elimination.

References

[1] "Digestion." https://en.wikipedia.org/wiki/Digestion. Accessed: April 2, 2015.

[2] "Digestive Disorders Health Center." http://www.webmd.com/digestive-disorders/the-lower-digestive-system. Accessed: April 2, 2015.

[3] A. Steffen, R. J. MacInnis, G. Joshy, G. G. Giles, E. Banks, and D. Roder, "Development and validation of a risk score predicting risk of colorectal cancer," *Cancer Epidemiol. Biomarkers Prev.*, vol. 23, pp. 2543–2552, Nov 2014. [DOI:10.1158/1055-9965.EPI-14-0206] [PubMed:25087576].

[4] "Jasmine." https://github.com/jasmine. Accessed: March 17, 2015.