

# **BREAST CANCER RISK ASSESSMENT TOOL**

By

**NUR AINI BINTI MOHD HAMID**

13563

**DISSERTATION**

Submitted to the Electrical & Electronics Engineering Programme  
in Partial Fulfillment of the Requirements  
for the Degree  
Bachelor of Engineering (Hons)  
(Electrical & Electronics Engineering)

MAY 2014

Universiti Teknologi PETRONAS  
Bandar Seri Iskandar  
31750 Tronoh  
Perak Darul Ridzuan

# **CERTIFICATION OF APPROVAL**

## **BREAST CANCER RISK ASSESSMENT TOOL**

by

Nur Aini binti Mohd Hamid

A project dissertation submitted to the  
Electrical & Electronics Engineering Programme  
Universiti Teknologi PETRONAS  
in partial fulfilment of the requirement for the  
Bachelor of Engineering (Hons)  
(Electrical & Electronics Engineering)

Approved by,

---

Assoc. Prof. Dr. Vijanth Segayan  
Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS  
TRONOH, PERAK

May 2014

## **CERTIFICATION OF ORIGINALITY**

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

---

Nur Aini binti Mohd Hamid

## **ABSTRACT**

Breast cancer is well known as one of the cancer that threatens woman. This disease threatens women, regardless race, ethnicity and demographics. Thus, in order to predict the risk of a woman being diagnosed with breast cancer, an individual risk assessment is developed. This is because early detection can help reduce the development of breast cancer. To date, few mathematical models have been proposed and developed for estimating the risk. The statistical model used is known as the Gail model where it provides the estimation of risk development of breast cancer over specific interval and period of time. The estimation mainly uses the women's demographic information as well as the risk factors as parameters. This model is suitable for woman, which is not suspected of having hereditary breast cancer. Original Gail model considering six attributes including age, age at first menstrual period, age at the time of the birth of first child, number of first-degree relatives that have breast cancer, breast biopsies and race/ethnicity. Using statistical modelling, the individual risk is calculated in terms of probabilities based on both the relative risk and the baseline hazard rate. In this project, the original risk factors are modified depending on the Malaysian demographic information to cater for Malaysia women population.

## **ACKNOWLEDGMENTS**

Special thanks to my supervisor Assoc. Prof. Dr. Vijanth Sagayan for his dedication helps to spend his precious time to teach, and guide me despite he has many other obligations. My gratitude towards Assoc. Prof. Dr Sarat, for sharing his expertise and teach me about statistical knowledge and always keeping his door open. Many thanks to my family back home for their sacrifices coupled with their continuous encouragement and support and heading me towards the stars. Special thanks to all the friends that help me completing this project, patiently teach me the basic language of the project and for together brainstorming in order to encounter problem faced. Each and every helps of them means the world to me.

## TABLE OF CONTENTS

LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
LIST OF ABBREVIATIONS .....	x
CHAPTER 1 INTRODUCTION .....	1
1.1 Background of Study.....	1
1.2 Problem Statement .....	3
1.3 Objectives.....	4
1.4 Scope of Study .....	4
CHAPTER 2 LITERATURE REVIEW AND THEORY .....	5
2.1 Gail Model.....	5
2.2 Case-Control vs Cohort Study.....	8
2.3 Statistical computing using survival analysis.....	10
2.4 Numerical computing technique .....	12
CHAPTER 3 METHODOLOGY AND PROJECT WORK.....	18
3.1 Procedure.....	18
3.2 Milestones .....	22
3.3 Gantt Chart .....	23
CHAPTER 4 RESULTS AND DISCUSSION .....	25
4.1 Project Progress .....	25
4.2 Assessment Coding .....	26
4.3 User interface .....	28
4.4 Results .....	30
CHAPTER 5 CONCLUSION AND RECOMMENDATION.....	35
5.1 Conclusion.....	35
5.2 Recommendation.....	35

REFERENCES.....	36
APPENDICES .....	38
Appendix A BCRA CLASS.....	39
Appendix B Converting CLASS .....	48
Appendix C RISK CLASS .....	54
Appendix D FORM 1 CODING .....	56
Appendix E BCRA CLASS .....	57

## LIST OF TABLES

<b>Table 1</b> Types of risk factors .....	2
<b>Table 2</b> Subscript definitions .....	12
<b>Table 3</b> Risk model and associated relative risks .....	13
<b>Table 4</b> Subscript definitions for logistic regression.....	14
<b>Table 5</b> Subscript definitions for linear interpolation.....	15
<b>Table 6</b> Projected probability based on years of follow up .....	16
<b>Table 7</b> Risk assessment gantt chart.....	23



## LIST OF FIGURES

<b>Figure 1</b> Mammography Result .....	2
<b>Figure 2</b> A graph of estimated baseline risk after 5 years follow up for original Gail model and the modified model .....	6
<b>Figure 3</b> Understanding case-control and cohort study.....	9
<b>Figure 4</b> Gail model method.....	18
<b>Figure 5</b> Flow chart of project.....	21
<b>Figure 6</b> Projected milestones .....	22
<b>Figure 7</b> Risk factors weightage ( $\beta$ ) .....	26
<b>Figure 8</b> Setup simple data .....	27
<b>Figure 9</b> Result of the calculation .....	27
<b>Figure 10</b> Absolute risk calculation .....	27
<b>Figure 11</b> Designing Assessment Form .....	28
<b>Figure 12</b> Main Form .....	29
<b>Figure 13</b> Assessment Form.....	29
<b>Figure 14</b> A copy of patient's record .....	30
<b>Figure 15</b> Age vs percentage risk.....	31
<b>Figure 16</b> Relationship between first menstrual age and risks.....	31
<b>Figure 17</b> Relationship between age of first live birth and risks.....	32
<b>Figure 18</b> Relationship between number of families with breast cancer and risks ....	33
<b>Figure 19</b> Relationship between number of family with breast cancer and risks .....	33

## **LIST OF ABBREVIATIONS**

**BCDDP**      Breast Cancer Detection Demonstration Project

**BCRA**      Breast Cancer Risk Assessment

**GUI**      Graphical User Interface

# CHAPTER 1

## INTRODUCTION

This chapter introduces and explains the project topic, “Breast Cancer Risk Assessment Tool”. A background study on this topic is highlighted followed by the problem statement, objectives to be achieved and finally the scope of study.

### 1.1 Background of Study

Statistics show that women are prone to be diagnosed with breast cancer than men. In fact, breast cancer is the second threat to women after lung cancer [1]. However, in Malaysia, breast cancer turns out to be the main cancer affecting women. Based on the report from National Cancer Registry, the Malaysian females are prone to develop the breast cancer case where one in twenty of them are at risk. Specifically, the risk is higher accordingly for the Malays, Indians and Chinese [2]. The actual cause of breast cancer remains unknown, but studies show many possible risks that might be associated with the breast cancer.

Current assessment tool use non-modifiable risk factors to diagnose risks. It also takes non-genetic risk factors into consideration based on lifestyle [11]. **Table 1** shows the different type of risk factors. By predicting the risk of having breast cancer, a woman can plan her screening frequency and type of risk-reduction program. Ministry of Health Malaysia had introduced breast cancer prevention and screening policy known as Breast Self-Examination (BSE), Clinical Breast Examination and Mammogram Screening [2]. Mammogram is divided into two types – screening and diagnostic. After the mammogram screening, further tests are conducted for confirmation of cancer through ultrasound test and later, the biopsy is done to remove cells. This is because not all breast cancers can be detected through a mammogram. **Figure 1** shows the example of mammogram results that is unclearly shows the sign of normal lump and cancer.

**Table 1** Types of risk factors

<b>Non-modifiable</b>	<b>Modifiable</b>
i. Gender and age	i. Breast feeding duration
ii. Individual cancer history	ii. Oral contraceptives
iii. Family cancer history	iii. Body mass index
iv. Genetics	iv. Dose-specific alcohol effect
v. Menstruation and menopause	v. Hormone replacement therapy (HRT)
vi. Breast density	vi. Pregnancy
	vii. Radiation exposure
	viii. Smoking
	ix. Dietary factors
	x. Physical activity
	xi. Antiperspirant



**Figure 1** Mammography Result

A few established mathematical models such as Gail, Claus, Ford and Tyrer-Cuzick (IBIS) are constructed based the breast cancer risk factors. The models' main purposes are to perform prediction of individual's unique breast cancer risk [5,6]. However, each model performs different predictions, depending on the type of cases. Thus, a few types of assessments or risk calculator had been developed based on each model. For instance, Gail model is suitable for woman, which is not suspected of having hereditary breast cancer. Meanwhile the Claus model only uses family history as the risk factor. The risk factor is then expanded into first and second degree relatives (including paternal line) with cancer. Last but not least, Tyrer-Cuzick model is much more suitable for hereditary cases which is for woman with a gene mutation [5,6]. Despite each model does have its own limitations, pros and cons, Gail model has been widely applied in determining the risk. This is because its reliability has been validated and can be adjusted for both non-modifiable and non-genetic risk factor. Thus, depending on the various types of research, the risk factor of the model has kept changing accordingly [6].

Each mathematical model is derived from several studies, which then will be tested and validated. However, to date, the Gail model is the only validated model which consists of logistic regressions, relative risks and odds ratios. This model eventually leads to an assessment tool that is widely used by health professionals and the National Cancer Institute (NCI) is known as Breast Cancer Risk Assessment (BCRA) Tool.

## **1.2 Problem Statement**

Current BCRA tools only use data from the White and African-American women and incorporates it with Gail model. Thus, the result from the assessment might be varied and not reliable to be applied by the Asians, especially Malaysians. Many studies have been conducted, but only within the region of Europe. It is believed that breast cancer cases are depending on demographic factors. Studies show higher incidence recorded in developed country compared to developing countries. Recently, the Asian finally comes out with their own cohort study. The Singaporeans have conducted a study on breast cancer screening project using the Gail model in 2012. However, they have not come out with their own assessment model yet. In October 2013, the

Koreans managed to come out with their very own Korean Risk Assessment Model for Breast Cancer Risk (KoBRAT), even though there is no programmed calculator yet for KoBRAT. The assessment comes out to be different from the original BCRA. They reasoned out that Gail model is overestimated, which beefs up the hypothesis that the original BCRA is not true for Asians. Nevertheless, both studies agree that the breast cancer risk is increase proportionally with age.

Therefore, in order to go in the Malaysia's model, the risk assessment model has to redevelop from scratch to earn it a user-friendly tool which able to tweak based on user demand. There is also a need to reduce and add another risk factor in this assessment to satisfy the Malaysian data (cancer and mortality data).

### **1.3 Objectives**

The objectives of this project are:

- To develop a user-friendly risk assessment software solutions for the perspective of Malaysian women.
- To evaluate the five year risk and lifetime breast cancer risk among women in Malaysia using developed software

### **1.4 Scope of Study**

The assessment tool requires an understanding in both structured and object oriented programming. Engineering students only learned structured programming which can only be used for the calculation of the risk. Thus, the program will be coded in C# language together with both types of program in order to create a good graphical user interface (GUI). This project also requires the development of the GUI for the ease of use. Thus, this user-friendly software is an extension of the National Cancer Institute (NCI) software which has been revised to fit the objective of this project. The study also requires an understanding of the risk factors and relative risks that affect the result of the assessment.

## CHAPTER 2

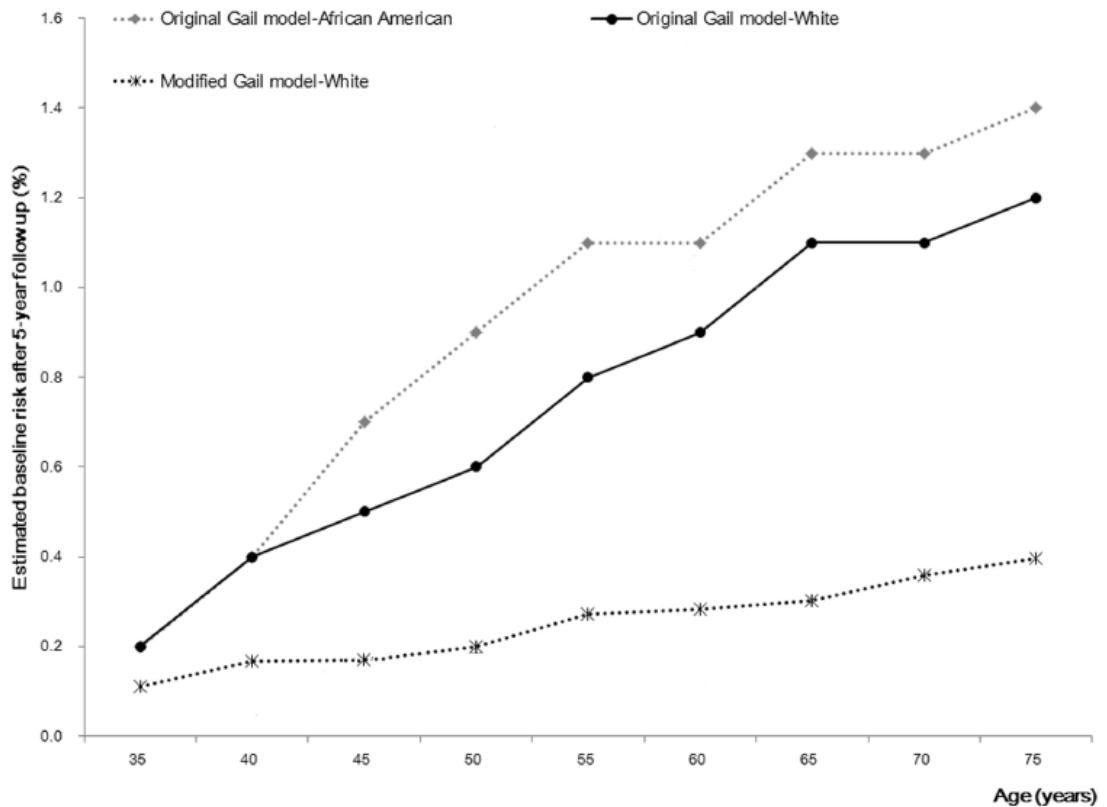
### LITERATURE REVIEW AND THEORY

This section reviews the critical points and theories covered in this project. It consists of the definition of Gail model and understanding the concept of cohort and case-control study. Next, statistical and mathematical approach is discussed in detail before integrating both approaches into engineering based programming.

#### 2.1 Gail Model

Dr Mitchell H. Gail (Gail et al, 1989), a Biostatistician along with his colleague develops the Gail model – a renowned research paper discussed in detail regarding the risk calculation using complex mathematical statistic to assess breast cancer [6]. The original Gail model or known as Gail Model 1, takes the Breast Cancer Detection Demonstration Project (BCDDP) data, a supervision program specially conducted for white females (the Caucasians) back on 1970s regarding breast cancer [7,9,10]. The information incorporates for Gail model regarding risk components are age of the woman, age at woman's first menstruation, age at woman's first live birth, number of first-degree relatives with breast cancer, number of prior breast biopsies, and presence of atypia on biopsy [7,11]. Later, ethnicity or race is being added into the risk factor as further studies conducted. Data from BCDDP is very significant towards contributing breast cancer research. It provides baseline age-specific breast cancer where an absolute risk can be estimated for breast cancers using the baseline value as shown in **Figure 2** [7]. BCDDP does not stop after the supervision program ended. Follow-up study was carried out which consists of four different phases in order to collect more data especially regarding mortality rate. Absolute risk projected is the probability of a woman developing breast cancer within a certain period of time. The absolute risk for a woman increases with age. Starting from age 30, thirteen age groups are being created with intervals of 5-years. By having intervals, the absolute risk can projected more accurately [7].

Recognized as Gail model 2, this modified model was used by statisticians from the National Surgical Adjuvant Breast and Bowel Project (NSABP) for determining eligibility for Breast Cancer Prevention Trial (BCPT) [14]. The NSABP main objective is to create projections for risk development of invasive breast cancer. This modified Gail model uses composite age-specific rates of invasive breast cancer for white woman and black women (which adjusted for the race part) that were derived from the Surveillance, Epidemiology, and End Results (SEER) data [7,9,14]. By using the modified Gail model also, the risk of breast cancer occurrence in the upcoming 5 years and throughout the lifetime was predicted for the participant of BCPT and validated using data from the placebo arm [6]. The result of the modified Gail model is certified as accurate only for woman who undergo mammograms annually. However, since mammogram is usually conducted by an older woman, the result from modified Gail model turns out to be overestimated for young women.



**Figure 2** A graph of estimated baseline risk after 5 years follow up for original Gail model and the modified model



As only Gail model has been validated, it is undoubtedly very useful and preferably chosen among the other models. The current BCRA tool available is built from the modified Gail model where it gives options for different race and ethnicity of the woman. The probabilities of developing breast cancer in woman are being calculated for 5 year projection and lifetime projection [7, 9, 10]. Despite the Gail model is widely used, it has its own limitation where it does not include second-degree relatives, the paternal side of family and age at diagnosis of the family.

This project uses modified Gail model as a risk assessment tool. As stated before, these are the key parameters for the modified Gail model.

- i. Age – older woman are prone to have breast cancer
- ii. Age at first menstrual period – woman that have early menstrual is said to have a higher chance of having breast cancer
- iii. Age at the time of the birth of first child (or has not given birth) – woman that gave birth at a young age are less likely to have breast cancer
- iv. Number of first-degree relatives with breast cancer – more relatives diagnosed with breast cancer increase the risk of having breast cancer
- v. Breast biopsies (number of biopsy and biopsy result) – the number of biopsies increases the risk of having cancer proportionally
- vi. Race/ethnicity – different racial group produces a different set of data

However, to fit in the Malaysian case study, more key parameters should be taken into considerations apart from the modified Gail parameters. Note that family history and genetics are two different types of unmodifiable risk factors. Genetics is more toward hereditary where a cancer patient carries a gene mutation or abnormal gene. This gene is passed from parents to the patient. Urbanization and current lifestyle might give impact which directly contributes to add the risk factors. These listed key parameters are among other risk factors that lead to the cancer. [5,9].

- i. Menopausal status
- ii. Duration of breast feeding
- iii. Oral contraceptives
- iv. Body mass index

- v. Mammography density
- vi. Dose-specific alcohol effect
- vii. Hormone replacement therapy (HRT)
- viii. Radiation exposure

These risk factors are categorized as non-genetic and modifiable risk factors. This is because; they are based on the woman's lifestyle. However, there are also some of the other factors that might contribute as risk factors such as antiperspirants and bras, but these factors have yet been used as one of the assessment risk factors.

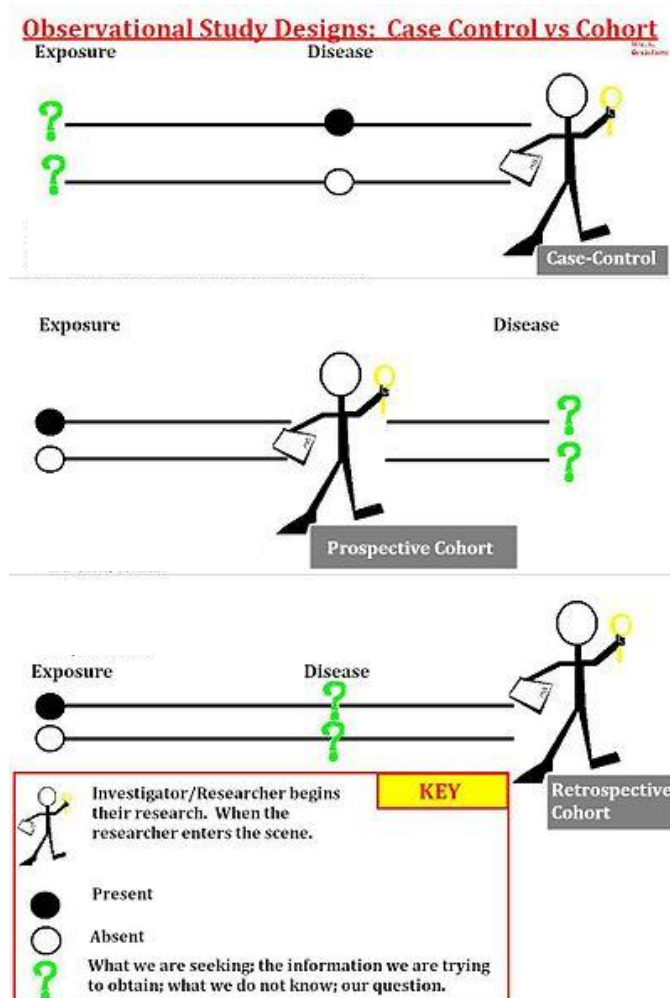
## **2.2 Case-Control vs Cohort Study**

The studies for cancer research involve two types of analysis – the analysis of case-control study and cohort study. A group of subjects that are chosen to compare either they are recognized as the disease carrier (the case) or non-disease carrier (the control) is defined as case-control study. These two groups are being observed on their exposure towards specific disease risk factors for validating the risk factors are the cause of disease. This retrospective study can be done to validate multiple risk factors at particular time [13]. Concurrently, cohort is the population of people, which undergo observation over the specific time. They are being observed on the exposure of possible variable that is being set at the beginning of the cohort study. After the given period, the results or outcome of them is recorded and will be compared with the other same group that is not being exposed from the possible variable [13].

In the early research on assessment, Gail model uses Caucasian woman (White woman) as the cohort for BCDDP [7]. The woman participated are mixture of age and different health background. By undergoing periodic examination; the participants have been under observation for 5 years. Next, from the result of the BCDDP after 5 years, a case-control study was started [9, 10]. The case are participants that receive any treatment regarding breast cancer either they are having the disease or potential to have the disease. This group of women will then be compared with the participants that did not have any signs or potential to be diagnosed so (known as the control). A list of risk factors that contributing towards

breast cancer development is acquired from the case-control study. This model was then repeated by using African American woman cohort [14].

To date, Asians have started to come out with their own cohort study. As different race or ethnic produce different estimation, the Koreans successfully came out with their KoBCRAT where it solely depends on their cancer centre cohorts. Again, from the cohort, they conduct the case-control study to smaller down the risk factors that lead to breast cancer [15]. The assessment still uses Gail model, however, with modification it becomes robust to be used for their own population. The study still maintains the division of two age's categorization – above and below fifty years old, but with additional and drop off risk factor. Singaporeans also started to conduct the screening project to predict the risk, especially for the older (50 years and above) using their own population (cohort) [16]. This screening project is an achievement for the Asians as it is being conducted other than European and American country.



**Figure 3** Understanding case-control and cohort study

### 2.3 Statistical computing using survival analysis

Survival analysis is defined as a compilation of statistical procedures for data analysis on the occurrence of events. The events can be disease incidence, injury, death, recovery, relapse from remission or any designated experience of interest. The objectives of this analysis are projecting time-to-event for a group of individuals, comparing time-to-events between two or more groups and to assess the relationship of co-variables to time-to-event [13].

The probability of developing breast cancer is made up of two components, mainly survival function  $S(t)$  and hazard function  $h(t)$  [7,13].

$$S(t) = e^{(-\int_0^t h(u) du)} \quad (1)$$

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{S(t)} \quad (2)$$

This two basic equations than being multiplied by each other to find the probability of developing of developing breast cancer between ages  $\alpha$  and  $\alpha + \tau$  as in equation 3. By adding age-dependent relative risk  $r(t)$  and age intervals, probability is determined by equation 4 and 5.

$$P\{\alpha, \tau, r(t)\} = \int_{\alpha}^{\alpha+\tau} h(t)S(t)dt \quad (3)$$

$$P\{\alpha, \tau, r(t)\} = \int_{\alpha}^{\alpha+\tau} h_1 r(t) e^{-\int_{\alpha}^t h_2 r(u) du} \left\{ \frac{S_2(t)}{S_2(\alpha)} \right\} dt \quad (4)$$

$$P\{\alpha, \tau, r(t)\} = \sum_j \left\{ \frac{h_{1j} r_j}{h_{1j} r_j + h_{2j}} \right\} \left\{ \frac{S_1(\tau_{j-1})}{S_1(\alpha)} \right\} \left\{ \frac{S_2(\tau_{j-1})}{S_2(\alpha)} \right\} \left\{ 1 - e^{-\Delta_j (h_{1j} r_j + h_{2j})} \right\} \quad (5)$$

where for equation 4 and 5,  $S_1$ ,  $S_2$ ,  $\tau_{j-1}$  and  $\tau_j$  are:

$$S_1(\tau_j) = S_1(\tau_{j-1})e^{(-h_{1j}r_j\Delta_j)} \quad (6)$$

$$S_2(t) = e^{(-\int_0^t h_2(u)du)} \quad (7)$$

$$\tau_{j-1} = \alpha ; \tau_j = \alpha + \tau \quad (8)$$

Baseline hazard,  $h_1(t)$  can be found out using the formula:

$$h_1(t) = h_1^*(t) \sum_{i=1}^I \frac{\rho_i(t)}{r_i(t)} = h^*(t)F(t) \quad (9)$$

To create the risk model and its associated relative risks, first we need to identify the unconditional logistic regression using the equation 10 and 11.

$$Odds = \frac{p}{1-p} = e^{\alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p} \quad (10)$$

$$logit p = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (11)$$

$\alpha$  represents the overall disease risk, while  $\beta_1$  is the fraction by which the risk is increased (or decreased) by every unit change in  $X_1$ ;  $\beta_2$  is the fraction by which the disease risk is altered by a unit change in  $X_2$  and so on. The logit of a proportion  $p$  is the logarithm of the corresponding odds. If an variable  $X$  has a coefficient  $\beta$ , then a unit increase in  $X$  increase the log odds by an amount equal to  $\beta$  [13].

The subscripts for the equations are defined in Table 2.

**Table 2** Subscript definitions

Subscript	Definition [3, 6, 13]
$S_1$	Probability of surviving the death due to breast cancer
$S_2$	Probability of surviving competing risks up to age $t$
$h_1$	Baseline age-specific hazard of developing cancer
$h_1^*$	Composite age-specific breast cancer incident rate
$h_2$	Risk of death due to other causes
$\rho_i(t)$	Proportion of woman of age $t$ in risk group $i$
$F(t)$	1 minus attributable risk fraction for age $t$
$r_j(t)$	Relative risk of the $i$ -th group compared to the baseline group
$\tau_j$	Time $j$ -th age interval
$\alpha$	Baseline age

#### 2.4 Numerical computing technique

The next step to create a risk assessment is by changing the statistical computing into numerical computing. Thus, from logit and odds formula in equation 10 and 11, the value of relative risk for each risk factor can be calculated as per equation 12.

$$rr = e^{\beta X} \quad (12)$$

Where  $X$  is the weights (0, 1 or 2) and  $\beta$  is the risk factors. The  $\beta$  value can be found from the logistic regressions express in equation 13. The risk factors together with its respective associated relative risks is shown in **Table 3**.

**Table 3** Risk model and associated relative risks

Risk factor		Associated relative risk
A. MENSTRUAL AGE		
≥14 (0)		1.000
12-13 (1)		1.099
<12 (2)		1.207
B. NO OF BIOPS		
Age < 50 years		
0 (0)		1.000
1 (1)		1.698
≥2 (2)		2.882
Age ≥ 50 years		
0 (0)		1.000
1 (1)		1.273
≥2 (2)		1.620
C. AGE OF FIRST LIFE BIRTH	NUMBER OF RELATIVES	
<20 (0)	0 (0)	1.000
	1 (1)	2.607
20-24 (1)	≥2 (2)	6.798
	0 (0)	1.244
	1 (1)	2.681
25-29 or no birth (2)	≥2 (2)	5.775
	0 (0)	1.548
	1 (1)	2.756
	≥2 (2)	4.907
≥30 (3)	0 (0)	1.927
	1 (1)	2.834
	≥2 (2)	4.169
D. BIOPSY WITH ATYPICAL HYPERPLASIA		
Not applicable or unknown (0)		1.000
No (1)		0.930
Yes (2)		1.820

$$\begin{aligned}
R = & -0.74948 + 0.09401 (m_a) + 0.52926 (n_b) + 0.21863 (f_b) \\
& + 0.95830 (n_r) + 0.01081 (a_c) - 0.28804 (n_b \times a_c) \\
& - 0.19081 (f_b \times n_r)
\end{aligned} \tag{13}$$

The subscripts for the equation 13 are defined in **Table 4**. All the subscript values are referred to **Table 3**.

**Table 4** Subscript definitions for logistic regression

Subscript	Definition [3, 6, 13]
$m_a$	Age at first menstruation
$n_b$	Number of biopsies
$f_b$	Age during first birth
$n_r$	Number of first degree relatives diagnosed with cancer
$a_c$	Age categories (<50 or $\geq$ 50)

Proceeding to the total relative risk calculation, it is calculated as :

$$rr_T = A \times B \times C \times D \tag{14}$$

After the calculation of total relative risk, some linear interpolation calculating was done with respect to years of follow up to calculate estimated risk, *ER*. It is basically using the initial relative risk ( $r_i$ ) and later relative risk ( $r_f$ ). The calculation of linear interpolation is as per equation 15. The estimated risk is calculated based on years of follow-up – 10, 20 and 30 years. However, if the years of follow up does not change the age group, the only initial calculation is required. For example, a woman aged 30 and requires a projected risk in 30 years. After 30 years, the woman change the age



categories to  $\geq 50$  years, which requires both initial relative risk (at age 30) and later relative risk (at age 60).

$$ER = r_{ie} + \frac{(V_2 - V_1)(r_i - r_{ie})}{rr_i - rr_{ie}} + \frac{(V_3 - V_1)(r_i - r_{ie})}{rr_i - rr_{ie}} \quad (15)$$

The subscripts for the equation 15 are defined in **Table 5**. The values for the subscripts are referring to **Table 6** which is the probabilities of breast cancer development based on follow up years,  $\tau$  for patients with age,  $\alpha$  and her respective relative risk value. The computations are then continue with a numerical integration technique which is based on trapezoidal rule. The full calculation is in the program as in **Appendix A - E**.

**Table 5** Subscript definitions for linear interpolation

Subscript	Definition
$V_1$	Initial risk
$V_2$	Calculated risk
$V_3$	Final risk
$r_l$	Initial projected probability
$r_{ie}$	Estimated initial projected probability
$r_i$	Final projected probability
$r_{ie}$	Estimated final projected probability
$rr_i$	Initial relative risk
$rr_{ie}$	Estimated initial relative risk
$rr_l$	Final relative risk

$rr_{ie}$	Estimated final relative risk
-----------	-------------------------------

**Table 6** Projected probability based on years of follow up

Initial age	Follow up years	Later relative risk	Initial relative risk						
			1.0	2.0	5.0	10.0	20.0	30.0	
20	10		0.0	0.1	0.2	0.5	1.0	1.4	
	20		0.5	1.0	2.5	4.9	9.5	14.0	
	30		1.7	3.4	8.3	15.9	29.3	40.5	
30	10		0.5	0.9	2.3	4.4	8.7	12.8	
	20		1.7	3.3	8.1	15.6	28.8	39.9	
	30	1.0	3.2	4.8	9.5	16.9	29.9	40.8	
		2.0	4.7	6.3	10.9	18.2	30.9	41.7	
		5.0	8.9	10.4	14.9	21.8	34.0	44.3	
		10.0	15.6	17.1	21.2	27.6	38.8	48.3	
		20.0	27.6	28.8	32.3	37.8	47.4	55.5	
		30.0	37.7	38.7	41.8	46.4	54.7	61.7	
	40	10		1.2	2.5	6.1	11.8	22.2	31.3
		20	1.0	2.8	4.0	7.5	13.1	23.4	32.4
		2.0	4.3	5.5	8.9	14.5	24.5	33.4	
		5.0	8.6	9.7	13.1	18.3	28.0	36.4	
		10.0	15.4	16.4	19.5	24.4	33.3	41.1	
		20.0	27.4	28.4	30.9	35.2	42.7	49.5	
		30.0	37.7	38.5	40.7	44.3	50.8	56.6	
30		1.0	4.4	5.6	9.1	14.6	24.6	33.5	
		2.0	7.4	8.6	11.9	17.3	27.0	35.6	
		5.0	15.9	17.0	20.0	24.9	33.7	41.5	
		10.0	28.3	29.2	31.8	35.9	43.4	50.0	
		20.0	47.5	48.1	50.0	53.1	58.5	63.4	
		30.0	61.2	61.6	63.1	65.3	69.3	72.8	

50	10	1.6	3.1	7.6	14.6	27.1	37.7
	20	3.2	6.4	15.1	27.9	47.8	61.9
	30	4.4	8.5	19.9	35.5	57.8	71.7
60	10	1.8	3.6	8.6	16.5	30.1	41.5
	20	3.0	5.9	14.0	25.9	44.6	58.2
70	10	1.4	2.7	6.7	12.9	24.1	33.7

## CHAPTER 3

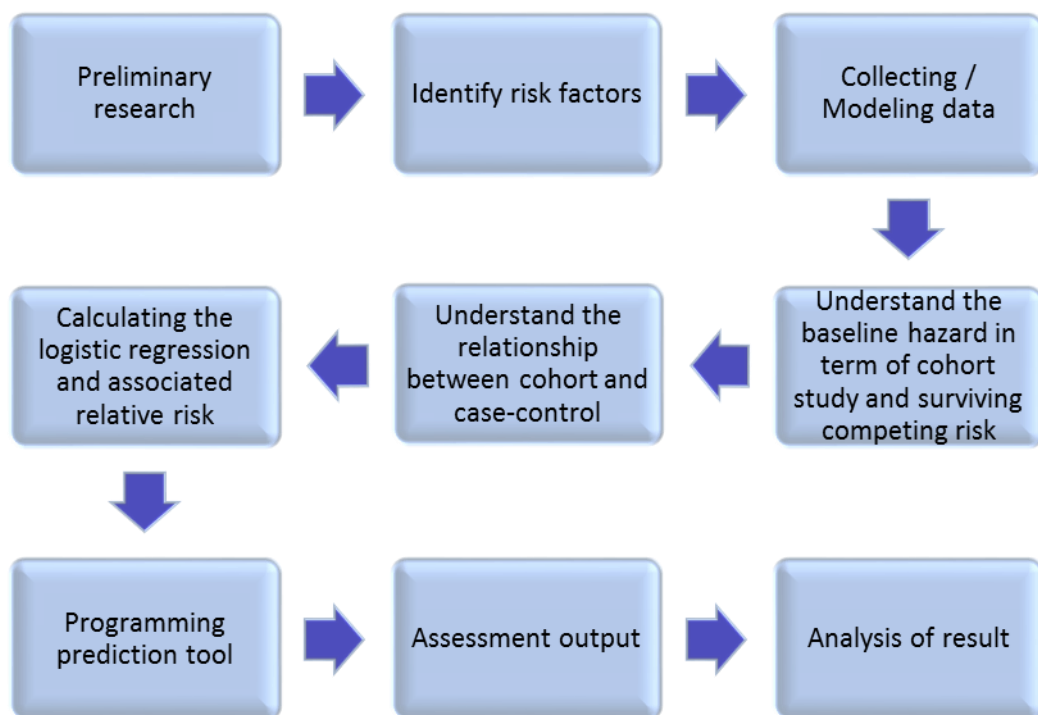
### METHODOLOGY AND PROJECT WORK

This chapter discusses the procedure and tools such as equipment and hardware to be implemented and utilized throughout the course of completing this project.

#### 3.1 Procedure

This project requires an understanding of the Gail model approach and the steps to program the collected data. The main output from this project is the projected risk of breast cancer in 5 years and throughout lifetime. The calculation is being done through a program in Microsoft Visual Studio using C# language.

##### 3.1.1 The Gail Model Approach



**Figure 4** Gail model method

**Preliminary research:** This is the first step to take in order to complete the project. The steps focused on collecting all relevant information that is related to the project. The information was obtained from trustworthy sources such as journals, articles, books, and technical papers. All the information obtained later will help in planning the best way to undertake the project and what to expect at the end of the project.

**Identify risk factors:** Anything that affects woman to be diagnosed as having breast cancer is known as risk factor. Identifying risk factors mean to decide whether to new risk factors or maintain the current risk factors, which is depending on mortality data as well as demographic.

**Collecting / Modelling data:** The data obtained (competing hazard and incidence hazard) should be transformed in the form of numerical data. Most data are given in term of numerical or graphical. The data that needs to be transformed into numerical data are age specific breast cancer composite index ( $h_1^*$ ) and age competing specific hazards ( $h_2$ ).

**Understanding the statistical equations:** The statistical equation listed in Chapter 2 consists of survival analysis, baseline hazard and age-specific hazard. Age and projected age are very significant for all the equations. Thus, in order to get more precise calculations, the age is being divided into 5 year interval starting from age 35 until 90.

**Calculating the regression and relative risk:** Logistic regression is the main component that is being used throughout the calculation. From the logit, a relative risk table can be produced. The total relative risk is then calculated by the multiplication of all relative risk from the list of risk factors.

**Programming prediction tool:**

Few steps of programming methodology have been planned and divided as below. Detail discussion regarding programming is shown in a flow chart in **Figure 5**.

i. Familiarization Process

This process is to familiarize with the software that will be used. The familiarization process includes learning C# language, learning to design user interface, and how object-oriented program works.

ii. User interface design

The reason for having a GUI is for users to understand what are the risk factors (parameters) that are contributing to breast cancer. It is also to ensure users can easily assess, understand and use to facilitate the actions. Some important elements that needed to be in the interface are input controls which consist of text fields, buttons, combo boxes, and date field. Other elements are navigational components and informational components such as message boxes.

iii. Converting user input

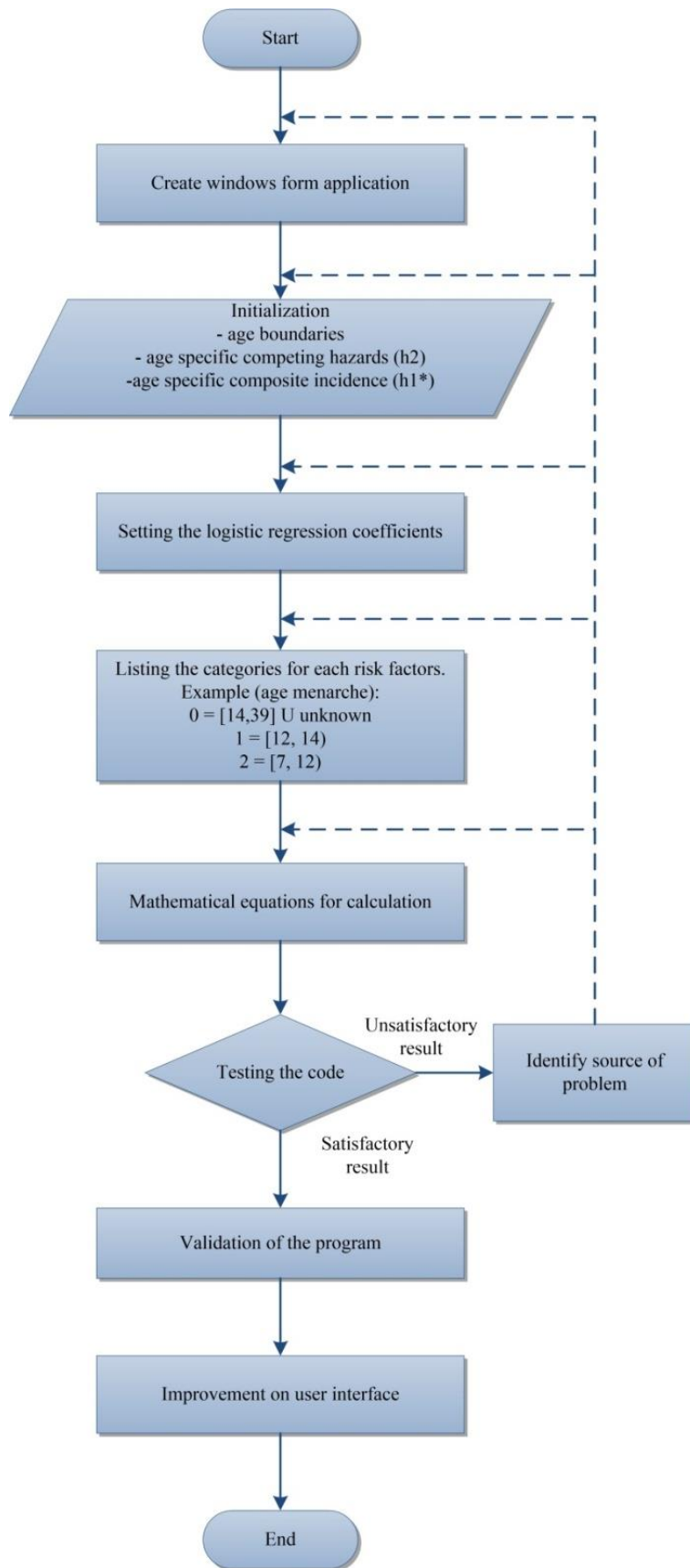
The coding does not read directly into the input value. Since the value taken from user consist a range of numbers, each input will be converted into the small scale value of numbers – 0, 1, 2 or 3. The reason changing into these small scale value is to ease the calculation part in the coding without using bigger memory.

iv. Testing the program and the improvement of the GUI

The program should be working properly by showing the outputs. Any errors and problems should be overcome. More elements can be added to GUI enhancement.

**Analysis of results:** After the data obtained and program shows output result, analysis of results will be done. All the results based on parameter manipulation will be used in order to justify and meet the objective of the project, which is to develop a user-friendly risk assessment software solutions for the perspective of Malaysian women

### 3.1.2 Calculation of predicting risk

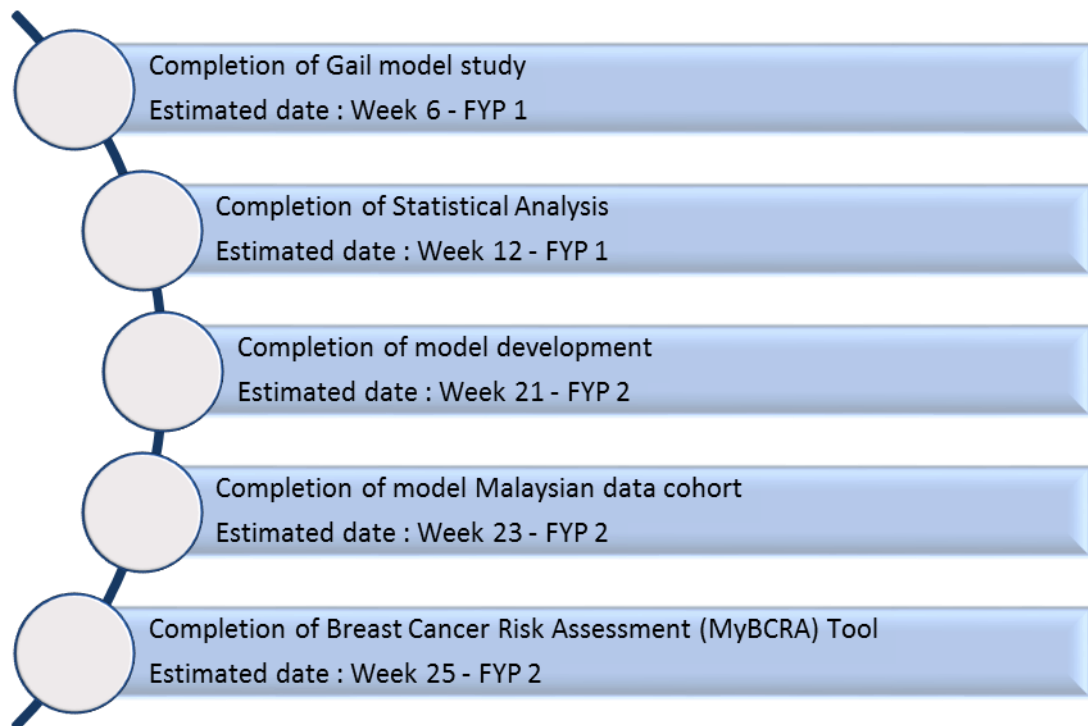


**Figure 5** Flow chart of project

This project is mainly on statistical and numerical computing programming. The software tool is based on C# object oriented programming using Visual Studio. The predicting tool is designed by creating user interface using the windows form application. The interface is designed in a form where designers can add controls to the form, set initial properties and write the handlers for specified event. The main program is programmed using class – a constructor that establishes the initial values, equations and calculation which act as a class. Several classes as per Appendix A-E are constructed to distinguish between the data initialization and calculation, converting input into a respected code value, and collecting input from the user. The predicting tool shall calculate 5 years projected risk as well as lifetime risk as the output.

### 3.2 Milestones

The target milestones for the project are shown in Figure 6



**Figure 6** Projected milestones







## **CHAPTER 4**

### **RESULTS AND DISCUSSION**

This chapter discusses the outcomes of the studies carried out to date, progress made and the steps to be taken in the project's progress towards its completion.

#### **4.1 Project Progress**

The programming has started the second part of this project. Due to complexity and analysis to understand the equation, the programming progress is being delayed. The initial plan on constructing the program is by using the reference from current BCRAT using SAS [11].

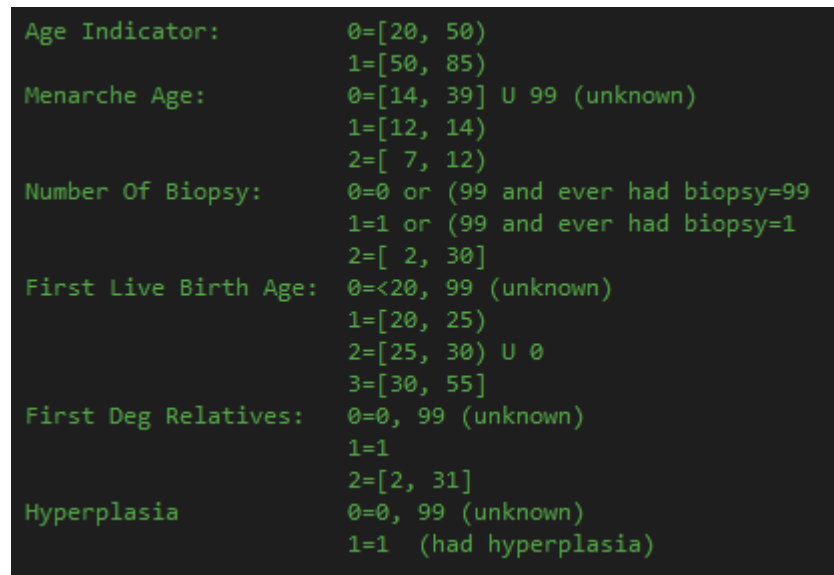
##### **4.1.1 Software Installation**

Up until week 14, the software chosen to code programming is MATLAB. However, Microsoft Visual Studio is chosen as finalized software as the code editor can be used in many languages – static and dynamic. Static programming language available is C#, C++ and VB.Net, while dynamic language available is IronPhyton, IronRuby and JavaScript. Languages such as HTML, XAML and SQL are also supported. For this project, code will be programmed in C# language. C# is another C style language. However, it takes time to understand the new object oriented language and class as well as familiarization with the applications – form application, console application and presentation application which lead with the delayed coding program. The software has to be installed twice due to Visual Studio is not supporting the three applications. It is meant for creating apps. Problem solved when the actual software to be used is Visual Studio for Desktop.

## 4.2 Assessment Coding

The whole project starts with transforming the statistical mathematics into numerical engineering method using trapezoidal rule and interpolation. The equations then being coded into a programming language. The assessment tool is also regarded as risk calculator where the coding results in projecting 5 year and lifetime risk. The coding use C# language (object oriented programming) using Visual Studio. The software used fits the objective of the project, which is to create a user friendly assessment tool.

The six parameters (risk factors) used in the program is not just a plain parameter. Every values that are chosen gives a significance value which contributes to its own weightage ( $\beta$ ) as stated in equation 11.



```
Age Indicator:      0=[20, 50)
                   1=[50, 85)
Menarche Age:      0=[14, 39] U 99 (unknown)
                   1=[12, 14)
                   2=[ 7, 12)
Number Of Biopsy:  0=0 or (99 and ever had biopsy=99
                   1=1 or (99 and ever had biopsy=1
                   2=[ 2, 30]
First Live Birth Age: 0=<20, 99 (unknown)
                   1=[20, 25)
                   2=[25, 30) U 0
                   3=[30, 55]
First Deg Relatives: 0=0, 99 (unknown)
                   1=1
                   2=[2, 31]
Hyperplasia        0=0, 99 (unknown)
                   1=1 (had hyperplasia)
```

**Figure 7** Risk factors weightage ( $\beta$ )

Using the situation as in Chapter 2, data are being set up as shown in **Figure 7** and shows the expected result.

```

// Setup Sample Data.

int currentAge = BcptConvert.GetCurrentAge(45);
int menarcheAge = BcptConvert.GetMenarcheAge("12");
int firstLiveBirthAge = BcptConvert.GetFirstLiveBirthAge("26");
int firstDegreeRel = BcptConvert.GetFirstDegRelatives("0");
int hadBiopsy = BcptConvert.GetEverHadBiopsy("0");
int numBiopsy = BcptConvert.GetNumberOfBiopsy("0");
int hyperPlasia = BcptConvert.GetHyperPlasia("0");
int race = BcptConvert.GetRace("1");

```

**Figure 8** Setup simple data

```

file:///C:/Users/USER/Documents...
5 year risk
This woman <age 45.00> = 0.90
Average woman <age 45.00> = 1.00

Lifetime risk
This woman <to age 90>: 10.6
Average woman <to age 90>: 11.9

```

**Figure 9** Result of the calculation

The data use of this program are logistic regression value, age category boundaries, age-specific competing hazard ( $h_2$ ), age specific composite incidence ( $h_1^*$ ), conversion factors and baseline hazards. **Appendix A** shows the data for this program. **Figure 10** shows one of the calculation in the program to find the absolute risk. The equation is derived from the equation 4 and 5 in Chapter 3. It takes the interval of 5 years and calculate based on average and absolute risk.

```

abs[i -1] =1.0-Math.Exp(-(rln[ni -1]*Math.Exp(sumbb[
i -1]) + rmu[ni -1]) * (ts - ti));

abs[i -1] = abs[i -1] * rln[ni -1] *Math.Exp(sumbb[
sumbb[i -1]) / (rln[ni -1] *Math.Exp(sumbb[
i -1]) + rmu[ni -1]);

```

**Figure 10** Absolute risk calculation

### 4.3 User interface

The Windows Form Application (WFA) in the Visual Studio turns out to be such a user-friendly user interface (UI) which suits the objective of this project. The form is a visual surface which displays information to the user and provides controls to the form – either to accept or display data. The figure shows an assessment form where the user will choose items from the combo box. They will then have the option to calculate their risk either in 5 years ahead or lifetime projection.

**Figure 11** Designing Assessment Form

The main form requires the user to key in first and last name as well as patient ID. Patient ID is required to save patient data so that a copy of the transcript of patients' risk can be monitored. This is part of user-friendly because it helps both patient and health professionals to save the patient's record. Indirectly, patients can understand and know what type of prevention method should she follow. Figure 12 shows the example of the main form of this project. Figure 13 shows the questionnaire that should be answered by the patient, which is the main calculation program is happening in this form.

**Breast Cancer Risk Assessment**

First Name: Nur Ani

Last Name: Mohd Hamid

Patient ID: BC 0001

Assessment Date: Monday, July 21, 2014

Enter

**Figure 12** Main Form

**Breast Cancer Risk Assessment**

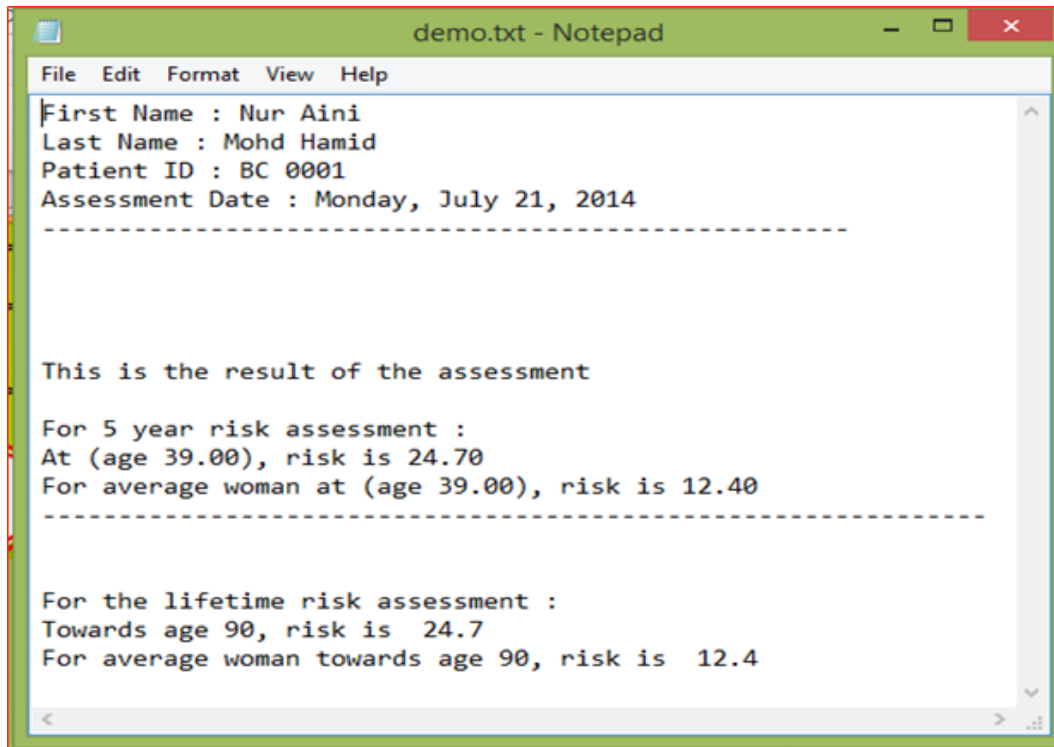
Please answer all the question. Then, click Calculate to get the risk assessment result

1. What is your current age? 39
2. When is your age during first menstrual period? 12
3. Age during first birth 26
4. Number of first degree relatives having breast cancer - mother, sisters, daughters 0
5. Have you ever had breast biopsy before? Yes
6. Number of breast biopsies undergone by you? 1
7. Has the biopsy turns out at least one as hyperlasia? No

Calculate Result Exit

**Figure 13** Assessment Form

The result of the assessment will be shown immediately and a record is save in .txt format according patients' ID as shown in Figure 14.



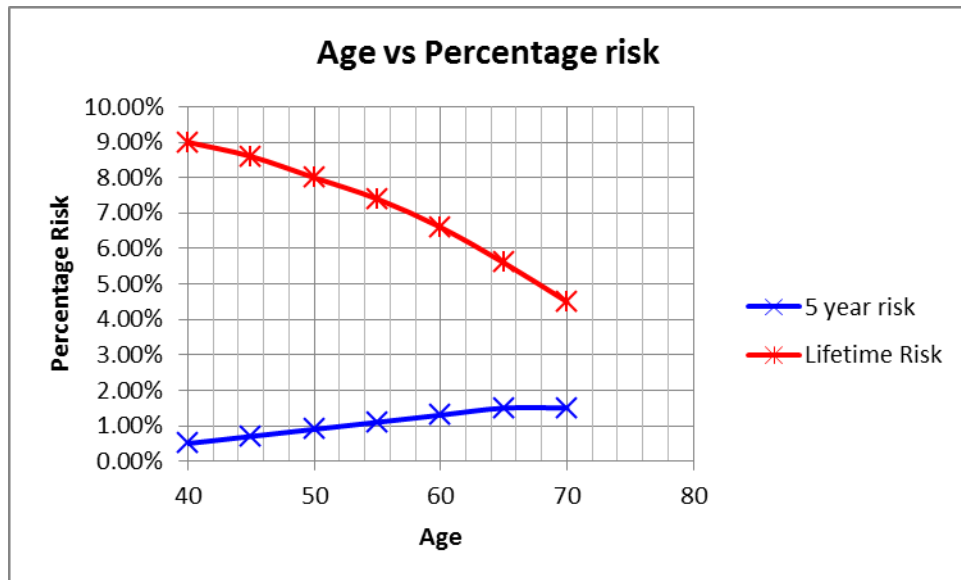
**Figure 14** A copy of patient's record

#### 4.4 Results

The results of the assessment are being discussed according to the listed risk factors. Note that for every risk factors, age and respected factors act as manipulated variable while the other factors remain constant.

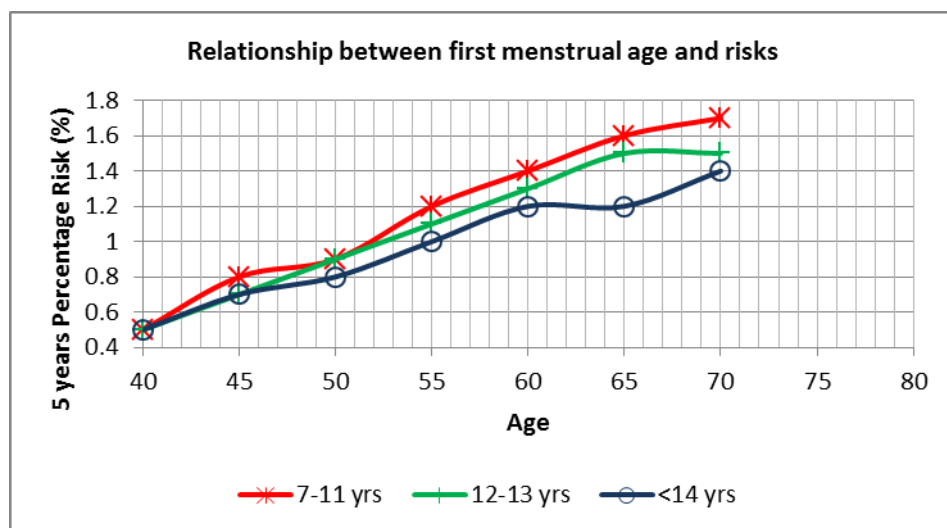
The first risk factor is age. As the age of a woman increase, their probability of having breast cancer is also increased. Figure 15 shows the result of the 5 year and lifetime risk calculation by changing the age and kept the other risk factors constant. From the figure, estimated 5 year risk of woman keep increasing based on every 5 year interval. However, the longer the projected year, the risk of having breast cancer is increasing because growth of cancer is a slow process. This means older women can be estimated as high risk of having breast cancer compared to the younger woman.





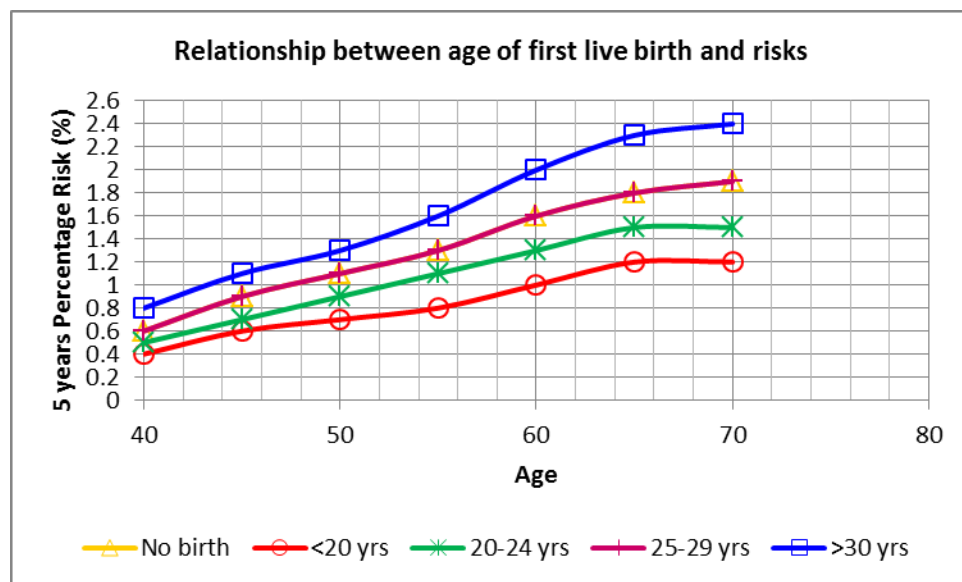
**Figure 15** Age vs percentage risk

Figure 16 shows the result of the 5 year percentage risk for menstrual age. The graph shows two significant data. The first one shows as age increase, the risk for any categories of menstrual age also increases. It clearly shows that early menstrual age contributes to higher risk compared to late menstrual cycle. The second one is for each age boundary, early menstruation clearly shows high risk. Therefore, the hypothesis are concluded as early menstrual age will contribute bigger chance in breast cancer development due to prolonged exposure to estrogen.



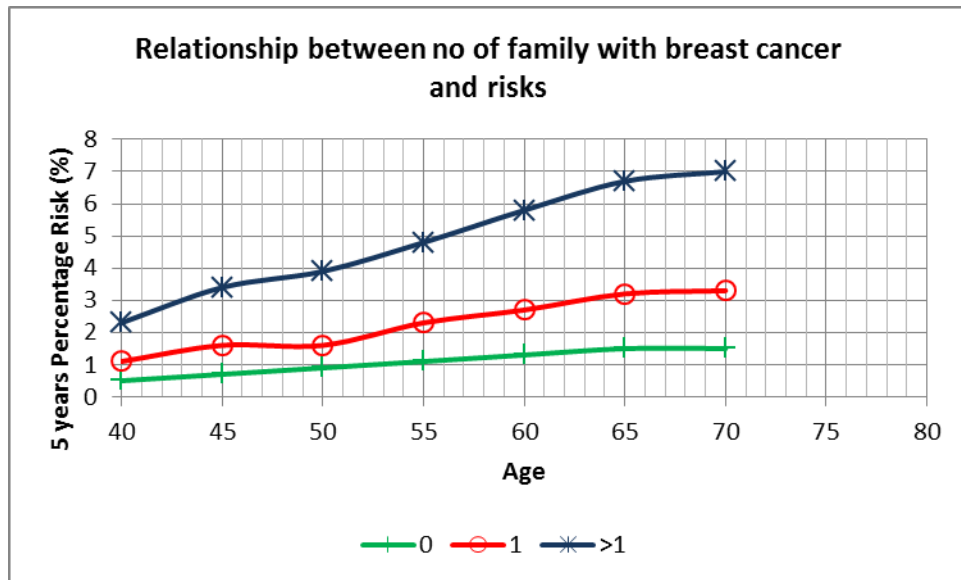
**Figure 16** Relationship between first menstrual age and risks

Figure 17 shows the result of the 5 year percentage risk for first birth age. The graph shows two significant data. The first one shows as age of women undergo assessment increase, the risk for any categories of first birth also increases. However, it still clearly shows that early first birth age contributes to smaller risk. This refers back to figure 15 where the risk is proportional to age. The second one is for each age boundary, early first birth age clearly shows lower risk. However, the risk values of no birth categories and 25-29 years boundary are same. Therefore, the hypothesis is concluded as late first birth age will contribute bigger chance in breast cancer development.



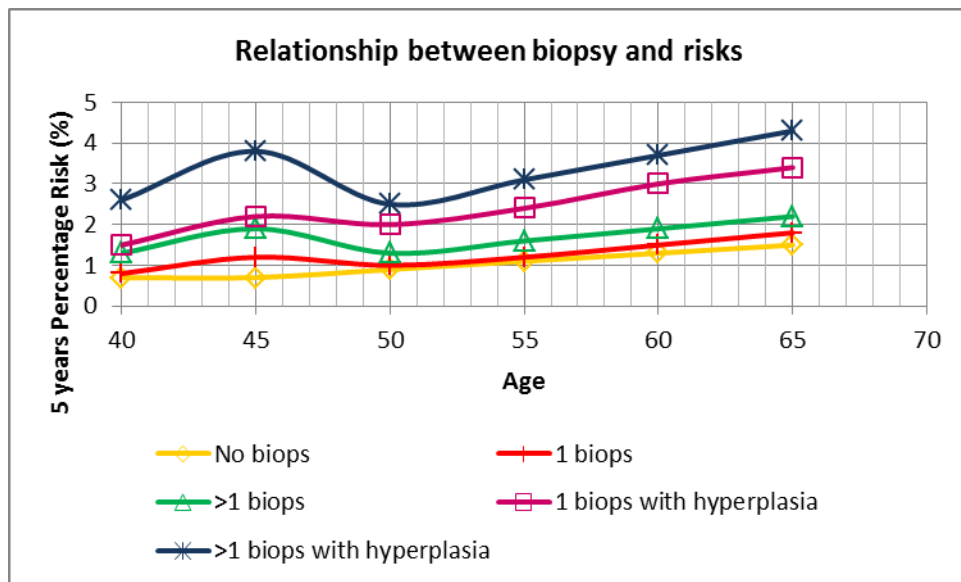
**Figure 17** Relationship between age of first live birth and risks

Figure 18 shows the result of the 5 year percentage risk for number of relatives diagnosed with breast cancer. The graph constantly shows for each age boundary, the risk is very high for those who have more than one relatives diagnosed with breast cancer. Unlike the other risks before, the highest risk can achieve up to almost 8%. The other two factors before does not exceed more than 3% before. This shows that the family history of a patient is very crucial in determining the risk.



**Figure 18** Relationship between number of families with breast cancer and risks

Figure 19 shows that women that never undergo any breast biopsies are less to be projected with breast cancer. However, any women with hyperplasia tends to have much higher risk, even though they only had undergone one biopsy.



**Figure 19** Relationship between number of family with breast cancer and risks

From the results of the programming, it is concluded that among the factors that contribute high percentage risk are:

- i. Older woman
- ii. Early menstrual age
- iii. Late giving birth
- iv. More family members diagnosed with breast cancer
- v. Undergo fews of breast biopsies diagnosed with hyperplasia

The results from the graphs shown the program is validated as per conclusion above. Indirectly, it shows that the program built are reliable and practical.

## **CHAPTER 5**

### **CONCLUSION AND RECOMMENDATION**

#### **5.1 Conclusion**

Although there is slight delayed in progress, the coding is workable and reliable. Problems have been overcome and the project continued. Detailed study carried out have provided a better understanding on the project, notably in areas programming language, GUI, and mathematical modelling. This project has been a learning experience to understand the application of object oriented programming in real world applications. This project also shows that each risk factor is interrelated and play an important role in determining the projected risk.

In conclusion, the progress made to date have been significant and will be essential towards the completion of the project. The success of this project will be a major improvement in the breast cancer assessment especially in Malaysian cohort.

#### **5.2 Recommendation**

Although the program is completed, it is better to explore additional risk factors, so that a much more reliable assessment can be obtained. However, in order to add more risk factors, the logistic regression equation should be changed. Further improvements may be required in order to obtain better results.

## REFERENCES

- [1] C. G. Varricchio and American Cancer Society Staff, "A Cancer Source Book for Nurses," 15 April 2004 8<sup>th</sup> ed. Sudbury, Massachusetts: Jones and Bartlett Learning, LLC, 2004, p580
- [2] M. Dahlui, S. Ramli and A. M. Bulgiba, "Breast Cancer Prevention and Control Programs in Malaysia," *Asia Pacific Journal of Cancer Prevention*, vol. 12, pp. 1-4, 2011.
- [3] L. Yuqin, C. Lihua, W. Xiaohai and C. Alan, "Implementation of Breast Cancer Risk Assessment Tool Using SAS®," in *PharmaSUG 2013 Conference Proceedings*, Chicago, pp. 1-8, 2013.
- [4] E. Amir, D. G. Evans, A. Shenton, F. Lalloo, A. Moran, C. Boggis, M. Wilson and A. Howell, "Evaluation of breast Cancer Risk Assessment Packages in the Family History Evaluation and Screening Programme," *Journal of Medical Genetics*, vol. 40, pp. 807-814, 26 June 2003.
- [5] D. Gareth, R. Evans and A. Howell "Breast Cancer Risk-assessment Models," *Breast Cancer Research*, vol. 9, pp. 213-220, 12 September 2007.
- [6] M. H. Gail, L. A. Brinton, D. P. Byar, D. K. Corle, S. B. Green, C. Schairer and J. J. Mulvihill, "Projecting Individualized Probabilities of Developing Breast Cancer For White Females Who are Being Examined Annually," *Journal of the National Cancer Institute*, vol. 81, pp. 1879-1886, 20 December 1989.
- [7] A. Decarli, S. Calza, G. Masala, C. Specchia, D. Palli and M. H. Gail, "Gail Model for Prediction of Absolute Risk of Invasive Breast Cancer: Independent Evaluation in the Florence-European Prospective Investigation Into Cancer and Nutrition Cohort," *Journal of the National Cancer Institute*, vol 98, issue 23, pp 1686-1693, 12 October 2006.
- [8] V. S. Pankratz, L. C. Hartmann, A. C. Degnim, R. A. Vierkant, K. Ghosh, C. M. Vachon, M. H. Frost, S. D. Maloney, C. Reynolds and J. C. Boughey, "Assessment of the Accuracy of the Gail Model in Women with Atypical Hyperplasia," *Journal of Clinical Oncology*, vol 26, issue 33, pp. 5374-5379, 20 November 2008.
- [9] P. Booyoung, M. Seung-hyun, S. Ae-sun, Chang. Myung-chul, C. Ji-yeob, K. Sungwan, H. Wonshik, N. Dong-young, A. Sei-hyun, K. Daehee, Y. Keun-young and P. K. Sue, "Korean Risk Assessment Model for Breast Cancer Risk Prediction", *PLOS ONE Journals*, pp. 1-9, 2013.

- [10] N.E. Breslow and N.E. Day, "Statistical Methods in Cancer Research, Vol 1 – The Analysis of Case-control Studies, " *International Agency for Research on Cancer (IARC) Scientific Publications*, pp. 5-338, 1980
- [11] C. W. Yee, O.W, Sze, T. P. Hoon, L. N. Q. Jie, H. G. Hui, W. C. Siong, C. K. Seng, C. K. Yew, T. M. Han and P. Ang, "Validation of the Gail model for Predicting Individual Breast Cancer Risk in Prospective Nationwide Study," *Breast Cancer Research*, vol. 14, pp. 1-12, 2012

## APPENDICES



## APPENDIX A

### BCRA CLASS

```
using System;
using System.Text;
using System.Xml;

namespace Assessment_Tool_2.Engine
{
    public class RiskCalculator
    {
        #region Members
        /** of covariate patterns in GAIL model
        const int NumCovPattInGailModel = 216;

        // matrix
        private double[,] bet2 = new double[8, 12];
        private double[] bet = new double[8];

        private double[] rf = new double[2];
        private double[] abs = new double[216];

        private double[] rlan = new double[14];
        private double[] rmu = new double[14];
        private double[] sumb = new double[216];
        private double[] sumbb = new double[216];
        private double[] t = new double[15];

        private double[,] rmu2 = new double[14, 12];
        private double[,] rlan2 = new double[14, 12];

        private double[,] rf2 = new double[2, 13];

        #endregion

        public RiskCalculator()
        {
            Initialize();
        }
        private void Initialize()
        {
            /* age categories boundaries */
            t[0] = 20.0;
            t[1] = 25.0;
            t[2] = 30.0;
            t[3] = 35.0;
            t[4] = 40.0;
            t[5] = 45.0;

            t[5] = 45.0;
            t[6] = 50.0;
            t[7] = 55.0;
            t[8] = 60.0;
            t[9] = 65.0;
            t[10] = 70.0;
            t[11] = 75.0;
            t[12] = 80.0;
            t[13] = 85.0;
            t[14] = 90.0;

            /* age specific competing hazards (h2) */
            rmu2[0, 0] = 49.3 * 0.00001; // [20:25)
            rmu2[1, 0] = 53.1 * 0.00001; // [25:30)
            rmu2[2, 0] = 62.5 * 0.00001; // [30:35)
            rmu2[3, 0] = 82.5 * 0.00001; // [35:40)
            rmu2[4, 0] = 130.7 * 0.00001; // [40:45)
            rmu2[5, 0] = 218.1 * 0.00001; // [45:50)
            rmu2[6, 0] = 365.5 * 0.00001; // [50:55)
            rmu2[7, 0] = 585.2 * 0.00001; // [55:60)
            rmu2[8, 0] = 943.9 * 0.00001; // [60:65)
            rmu2[9, 0] = 1502.8 * 0.00001; // [65:70)
            rmu2[10, 0] = 2383.9 * 0.00001; // [70:75)
        }
    }
}
```

```

rmu2[11, 0] = 3883.2 * 0.00001; // [75:80]
rmu2[12, 0] = 6682.8 * 0.00001; // [80:85]
rmu2[13, 0] = 14490.8 * 0.00001; // [85:90]

/* age specific competing hazards (h2) for "average woman" */
rmu2[0, 1] = 44.12 * 0.00001; // [20,25]
rmu2[1, 1] = 52.54 * 0.00001; // [24,30]
rmu2[2, 1] = 67.46 * 0.00001; // [30,35]
rmu2[3, 1] = 90.92 * 0.00001; // [34,40]
rmu2[4, 1] = 125.34 * 0.00001; // [40,45]
rmu2[5, 1] = 195.70 * 0.00001; // [44,50]
rmu2[6, 1] = 329.84 * 0.00001; // [50,55]
rmu2[7, 1] = 546.22 * 0.00001; // [54,60]
rmu2[8, 1] = 910.35 * 0.00001; // [60,65]
rmu2[9, 1] = 1418.54 * 0.00001; // [64,70]
rmu2[10, 1] = 2259.35 * 0.00001; // [70,75]
rmu2[11, 1] = 3611.46 * 0.00001; // [74,80]
rmu2[12, 1] = 6136.26 * 0.00001; // [80,84]
rmu2[13, 1] = 14206.63 * 0.00001; // [84,90]

/* age specific breast cancer composite incidence (h1*) */

rlan2[0, 0] = 1.0 * 0.00001; // [20:25]
rlan2[1, 0] = 7.6 * 0.00001; // [25:30]
rlan2[2, 0] = 26.6 * 0.00001; // [30:35]
rlan2[3, 0] = 66.1 * 0.00001; // [35:40]
rlan2[4, 0] = 126.5 * 0.00001; // [40:45]
rlan2[5, 0] = 186.6 * 0.00001; // [45:50]
rlan2[6, 0] = 221.1 * 0.00001; // [50:55]
rlan2[7, 0] = 272.1 * 0.00001; // [55:60]
rlan2[8, 0] = 334.8 * 0.00001; // [60:65]
rlan2[9, 0] = 392.3 * 0.00001; // [65:70]
rlan2[10, 0] = 417.8 * 0.00001; // [70:75]
rlan2[11, 0] = 443.9 * 0.00001; // [75:80]
rlan2[12, 0] = 442.1 * 0.00001; // [80:85]
rlan2[13, 0] = 410.9 * 0.00001; // [85:90]

/* age specific breast cancer composite incidence (h1*)-"average
woman" */
rlan2[0, 1] = 1.22 * 0.00001; // [20:25]
rlan2[1, 1] = 7.41 * 0.00001; // [25:30]
rlan2[2, 1] = 22.97 * 0.00001; // [30:35]
rlan2[3, 1] = 56.49 * 0.00001; // [35:40]
rlan2[4, 1] = 116.45 * 0.00001; // [40:45]
rlan2[5, 1] = 195.25 * 0.00001; // [45:50]
rlan2[6, 1] = 261.54 * 0.00001; // [50:55]
rlan2[7, 1] = 302.79 * 0.00001; // [55:60]
rlan2[8, 1] = 367.57 * 0.00001; // [60:65]
rlan2[9, 1] = 420.29 * 0.00001; // [65:70]
rlan2[10, 1] = 473.08 * 0.00001; // [70:75]
rlan2[11, 1] = 494.25 * 0.00001; // [75:80]
rlan2[12, 3] = 479.76 * 0.00001; // [80:85]
rlan2[13, 3] = 401.06 * 0.00001; // [85:90]

/* age specific breast cancer composite incidence (h1*)*/
rlan2[0, 4] = 0.00002696; // [20:25]
rlan2[1, 4] = 0.00011295; // [25:30]
rlan2[2, 4] = 0.00031094; // [30:35]
rlan2[3, 4] = 0.00067639; // [35:40]
rlan2[4, 4] = 0.00119444; // [40:45]
rlan2[5, 4] = 0.00187394; // [45:50]
rlan2[6, 4] = 0.00241504; // [50:55]
rlan2[7, 4] = 0.00291112; // [55:60]
rlan2[8, 4] = 0.00310127; // [60:65]
rlan2[9, 4] = 0.00366560; // [65:70]
rlan2[10, 4] = 0.00393132; // [70:75]
rlan2[11, 4] = 0.00408951; // [75:80]
rlan2[12, 4] = 0.00396793; // [80:85]

```

```

rlan2[13, 4] = 0.00363712;           // [85:90)

rlan2[0, 5] = 2.00 * 0.00001;       // [20:25)
rlan2[1, 5] = 7.10 * 0.00001;       // [25:30)
rlan2[2, 5] = 19.70 * 0.00001;      // [30:35)
rlan2[3, 5] = 43.80 * 0.00001;      // [35:40)
rlan2[4, 5] = 81.10 * 0.00001;      // [40:45)
rlan2[5, 5] = 130.70 * 0.00001;     // [45:50)
rlan2[6, 5] = 157.40 * 0.00001;     // [50:55)
rlan2[7, 5] = 185.70 * 0.00001;     // [55:60)
rlan2[8, 5] = 215.10 * 0.00001;     // [60:65)
rlan2[9, 5] = 251.20 * 0.00001;     // [65:70)
rlan2[10, 5] = 284.60 * 0.00001;    // [70:75)
rlan2[11, 5] = 275.70 * 0.00001;    // [75:80)
rlan2[12, 5] = 252.30 * 0.00001;    // [80:85)
rlan2[13, 5] = 203.90 * 0.00001;    // [85:90)

// logistic regression coefficients - GAIL model (BCDDP)

bet2[0, 0] = -0.7494824600;           // intercept
bet2[1, 0] = 0.0108080720;           // age >= 50 indicator
bet2[2, 0] = 0.0940103059;           // age first menstrual
bet2[3, 0] = 0.5292641686;           // # of breast biopsy
bet2[4, 0] = 0.2186262218;           // age 1st live birth
bet2[5, 0] = 0.9583027845;           // relatives
bet2[6, 0] = -0.2880424830;
bet2[7, 0] = -0.1908113865;

/* age 1st live birth * # 1st degree rel */
// conversion factors (1-attributable risk) used in BCPT model

rf2[0, 0] = 0.5788413;               // age < 50
rf2[1, 0] = 0.5788413;               // age >=50

// conversion factors (1-attributable risk) used for "avg woman"

rf2[0, 3] = 1.0;                     // age < 50
rf2[1, 3] = 1.0;                     // age >=50

public double CalculateAbsoluteRisk(
    int CurrentAge
    , int ProjectionAge
    , int AgeIndicator
    , int NumberOfBiopsy
    , int MenarcheAge
    , int FirstLiveBirthAge
    , int FirstDegRelatives
    , int EverHadBiopsy
    , int ihyp
    , double rhyp
    , int irace
)
{
    return CalculateRisk(1
        , CurrentAge
        , ProjectionAge
        , AgeIndicator
        , NumberOfBiopsy
        , MenarcheAge
        , FirstLiveBirthAge
        , EverHadBiopsy
        , FirstDegRelatives
        , ihyp
        , rhyp
        , irace
    );
}
public double CalculateAverageRisk(

```

```

    int CurrentAge
    , int ProjectionAge
    , int AgeIndicator
    , int NumberOfBiopsy
    , int MenarcheAge
    , int FirstLiveBirthAge
    , int FirstDegRelatives
    , int EverHadBiopsy
    , int ihyp
    , double rhyp
    , int irace
    )
{
    return CalculateRisk(2
        , CurrentAge
        , ProjectionAge
        , AgeIndicator
        , NumberOfBiopsy
        , MenarcheAge
        , FirstLiveBirthAge
        , EverHadBiopsy
        , FirstDegRelatives
        , ihyp
        , rhyp
        , irace
        );
}
private double CalculateRisk(
    int riskindex
    , int CurrentAge
    , int ProjectionAge
    , int AgeIndicator
    , int NumberOfBiopsy
    , int MenarcheAge
    , int FirstLiveBirthAge
    , int EverHadBiopsy
    , int FirstDegRelatives
    , int ihyp
    , double rhyp
    , int irace)
{
    double retval = 0.0;
    int i, j, k, n, ni, ns, incr, ilev;
    double r, ti, ts, abss = 0.0;
    double[,] rSiTox2 = new double[216, 9];
    n = 216; /* ** age categories boundaries */
    r = 0.0;
    ni = 0;
    ns = 0;
    ti = (double)(CurrentAge);
    ts = (double)(ProjectionAge);

    /*setting BETA to race specific lnRR*/
    for (i = 0; i < 8; i++)
    { bet[i] = bet2[i, irace - 1]; }

    if (irace == 2)
    {
        if (MenarcheAge == 2)
        {
            MenarcheAge = 1;
            FirstLiveBirthAge = 0;
        }
    }

    for (i = 1; i <= 15; ++i)
    {
        if (ti < t[i - 1])
        {
            ni = i - 1; /* ni holds the index for current */
            break;
        }
    }
}

```

```

for (i = 1; i <= 15; ++i)
{
    if (ts <= t[i - 1])
    {
        ns = i - 1; /* ns holds the index for risk as */
        break;
    }
}
incr = 0;
if (riskindex == 2 && irace < 7)
{
    incr = 3;
}

int cindx = 0; //column index
cindx = incr + irace - 1;

for (i = 0; i < 14; ++i)
{
    rmu[i] = rmu2[i, cindx];
    rlan[i] = rlan2[i, cindx];
}

rf[0] = rf2[0, incr + irace - 1];
rf[1] = rf2[1, incr + irace - 1];
if (riskindex == 2 && irace >= 7)
{
    rf[0] = rf2[0, 12];
    rf[1] = rf2[1, 12];
}

if (riskindex >= 2) {
    /* set risk factors to */
    MenarcheAge = 0; // baseline age menarchy
    NumberOfBiopsy = 0; // # of previous biop
    FirstLiveBirthAge = 0; // age 1st live birth
    FirstDegRelatives = 0; // # 1st degree relat
    rhyp = 1.0; // set hyperplasia to 1.0
}

ilev = AgeIndicator * 108 + MenarcheAge * 36 +
    NumberOfBiopsy * 12 + FirstLiveBirthAge * 3 +
    FirstDegRelatives + 1; /* matrix of */

for (k = 0; k < 216; ++k)
{
    /* col1: intercept 0 */
    r8iTox2[k, 0] = 1.0;
}
for (k = 0; k < 108; ++k)
{
    /* col2: indicator for age */
    r8iTox2[k, 1] = 0.0;
    r8iTox2[108 + k, 1] = 1.0;
}
for (j = 1; j <= 2; ++j)
{
    /* col3: age menarchy cate */
    for (k = 1; k <= 36; ++k)
    {
        r8iTox2[(j - 1)*108 + k - 1, 2] = 0.0;

        r8iTox2[(j - 1)*108 + 36 + k - 1, 2] = 1.0;
        r8iTox2[(j - 1)*108 + 72 + k - 1, 2] = 2.0;
    }
}
for (j = 1; j <= 6; ++j)
{
    /* col4: # biopsy cate */
    for (k = 1; k <= 12; ++k)
    {
        r8iTox2[(j - 1)* 36 + k - 1, 3] = 0.0;
        r8iTox2[(j - 1)* 36 + 12 + k - 1, 3] = 1.0;
        r8iTox2[(j - 1)* 36 + 24 + k - 1, 3] = 2.0;
    }
}
}

```

```

for (j = 1; j <= 18; ++j)
{
  /* col5: age 1st live birt */
  for (k = 1; k <= 3; ++k)
  {
    r8iTox2[(j - 1) * 12 + k - 1, 4] = 0.0;
    r8iTox2[(j - 1) * 12 + 3 + k - 1, 4] = 1.0;
    r8iTox2[(j - 1) * 12 + 6 + k - 1, 4] = 2.0;
    r8iTox2[(j - 1) * 12 + 9 + k - 1, 4] = 3.0;
  }
}
for (j = 1; j <= 72; ++j)
{
  /* col6: # 1st degree re */
  r8iTox2[(j - 1) * 3 + 1 - 1, 5] = 0.0;
  r8iTox2[(j - 1) * 3 + 2 - 1, 5] = 1.0;
  r8iTox2[(j - 1) * 3 + 3 - 1, 5] = 2.0;
}
for (i = 0; i < 216; ++i)
{
  /* col8: age 1st live*# r */
  /* col7: age*#biop intera */
  r8iTox2[i, 6] = r8iTox2[i, 1] * r8iTox2[i, 3];
  r8iTox2[i, 7] = r8iTox2[i, 4] * r8iTox2[i, 5];
}
for (i = 0; i < 216; ++i)
{
  r8iTox2[i, 8] = 1.0;
}

/* ** Computation of breast cancer risk */
/* ** sum(bi*Xi) for all covariate patterns */
for (i = 0; i < 216; ++i)
{
  /* ln relative risk from BCDDP */
  sumb[i] = 0.0;
  for (j = 0; j < 8; ++j)
  {
    sumb[i] += bet[j] * r8iTox2[i, j];
  }
}

for (i = 1; i <= 108; ++i)
{
  /* eliminate int */
  sumbb[i - 1] = sumb[i - 1] - bet[0];
}
for (i = 109; i <= 216; ++i)
{
  /* eliminate intercept */
  sumbb[i - 1] = sumb[i - 1] - bet[0] - bet[1];
}
for (j = 1; j <= 6; ++j)
{
  /* age specific baseline hazard */

  rlan[j - 1] *= rf[0];
}
for (j = 7; j <= 14; ++j)
{
  /* age specific baseline hazard a */
  rlan[j - 1] *= rf[1];
}
i = ilev; /* index ilev of range 1- */
/* setting i to covariate p */
sumbb[i - 1] += Math.Log(rhyp);
if (i <= 108)
{
  sumbb[i + 107] += Math.Log(rhyp);
}

```

```

if (ts <= t[ni])
{
    /* same 5 year age risk in */
    /* age & projection age wi */
    abs[i - 1] = 1.0 - Math.Exp(-(r1an[ni - 1] *
Math.Exp(sumbb[i - 1]) + rmu[ni - 1]) * (ts - ti));
    abs[i - 1] = abs[i - 1] * r1an[ni - 1] *
Math.Exp(sumbb[i - 1]) / (r1an[ni - 1] *
Math.Exp(sumbb[i - 1]) + rmu[ni - 1]);/* breast cancer */
}
else
{
    /* 5 year age risk interval */
    /* calculate risk from */
    /* 1st age interval */
    /* age & projection age not i */
    abs[i - 1] = 1.0 - Math.Exp(-(r1an[ni - 1] * Math.Exp(sumbb[i - 1])
+ rmu[ni - 1]) * (t[ni] - ti));
    abs[i - 1] = abs[i - 1] * r1an[ni - 1] * Math.Exp(sumbb[i - 1]) /
(r1an[ni - 1]
        * Math.Exp(sumbb[i - 1]) + rmu[ni - 1]); /* age in
*/
    /* risk f */
    if (ns - ni > 0)
    {
        if ((double)(ProjectionAge) > 50.0 && (double)(CurrentAge) <
50.0)
        {
            r = 1.0 - Math.Exp(-(r1an[ns - 1] * Math.Exp(sumbb[i + 107])
+ rmu[ns - 1]) * (ts - t[ns - 1]));
            n = r * r1an[ns - 1] * Math.Exp(sumbb[i + 107]) / (r1an[ns -
1] * Math.Exp(sumbb[i + 107]) + rmu[ns - 1]);
            r *= Math.Exp(-(r1an[ni - 1] * Math.Exp(sumbb[i - 1]) +
rmu[ni - 1]) * (t[ni] - ti));

            if (ns - ni > 1)
            {
                MenarcheAge = ns - 1;
                for (j = ni + 1; j <= MenarcheAge; ++j)
                {
                    if (t[j - 1] >= 50.0)
                    {
                        n *= Math.Exp(-(r1an[j - 1] * Math.Exp(sumbb[
i + 107]) + rmu[j - 1]) * (t[
j] - t[j - 1]));
                    }
                    else
                    {
                        n *= Math.Exp(-(r1an[j - 1] * Math.Exp(sumbb[
i - 1]) + rmu[j - 1]) * (t[j]
- t[j - 1]));
                    }
                }
            }
            abs[i - 1] += n;
        }
    }
    else
    {
        /* calculate risk from */
        /* last age interval */
        /* ages do not stradle */
        n = 1.0 - Math.Exp(-(r1an[ns - 1] * Math.Exp(sumbb[i - 1])
+ rmu[ns - 1]) * (ts - t[ns - 1]));
        n = n * r1an[ns - 1] * Math.Exp(sumbb[i - 1]) / (
            r1an[ns - 1] * Math.Exp(sumbb[i - 1]) +
            rmu[ns - 1]);
        n *= Math.Exp(-(r1an[ni - 1] * Math.Exp(sumbb[i - 1]) +
            rmu[ni - 1]) * (t[ni] - ti));
    }
}
}

```

```

        if (ns - ni > 1)
        {
            MenarcheAge = ns - 1;
            for (j = ni + 1; j <= MenarcheAge; ++j)
            {
                r *= Math.Exp(-(rlan[j - 1] * Math.Exp(sumbb[i - 1]) +
                    rmu[j - 1]) * (t[j] - t[j - 1]));
            }
            abs[i - 1] += r;
        }
    }
    if (ns - ni > 1)
    {
        if ((double)(ProjectionAge) > 50.0 && (double)(CurrentAge) <
50.0)
        {
            /* calculate risk from */
            /* intervening age int */
            MenarcheAge = ns - 1;
            for (k = ni + 1; k <= MenarcheAge; ++k)
            {
                if (t[k - 1] >= 50.0)
                {
                    r = 1.0 - Math.Exp(-(rlan[k - 1] * Math.Exp(sumbb[
                        i + 107]) + rmu[k - 1]) * (t[k] -
                        t[k - 1]));
                    r = r * rlan[k - 1] * Math.Exp(sumbb[i +
                        107]) / (rlan[k - 1] * Math.Exp(sumbb[
                        i + 107]) + rmu[k - 1]);
                }
                else
                {
                    r = 1.0 - Math.Exp(-(rlan[k - 1] * Math.Exp(sumbb[
                        i - 1]) + rmu[k - 1]) * (t[k] -
                        t[k - 1]));
                    r = r * rlan[k - 1] * Math.Exp(sumbb[i - 1]
                        ) / (rlan[k - 1] * Math.Exp(sumbb[i -
                        1]) + rmu[k - 1]);
                }
            }
            r *= Math.Exp(-(rlan[ni - 1] * Math.Exp(sumbb[i - 1])
                + rmu[ni - 1]) * (t[ni] - ti));
            NumberOfBiopsy = k - 1;
            for (j = ni + 1; j <= NumberOfBiopsy; ++j)
            {
                if (t[j - 1] >= 50.0)
                {
                    r *= Math.Exp(-(rlan[j - 1] * Math.Exp(sumbb[
                        i + 107]) + rmu[j - 1]) * (t[
                        j] - t[j - 1]));
                }
                else
                {
                    r *= Math.Exp(-(rlan[j - 1] * Math.Exp(sumbb[
                        i - 1]) + rmu[j - 1]) * (t[j]
                        - t[j - 1]));
                }
            }
            abs[i - 1] += r;
        }
    }
    else
    {
        /* calculate risk from */
        /* intervening age int */
        MenarcheAge = ns - 1;
        for (k = ni + 1; k <= MenarcheAge; ++k)
        {

```



```

        r = 1.0 - Math.Exp(-(rlan[k - 1] * Math.Exp(sumbb[i - 1] +
        1)) + rmu[k - 1]) * (t[k] - t[k - 1]));
        r = r * rlan[k - 1] * Math.Exp(sumbb[i - 1]) /
        (rlan[k - 1] * Math.Exp(sumbb[i - 1]) +
        rmu[k - 1]);
        r *= Math.Exp(-(rlan[ni - 1] * Math.Exp(sumbb[i - 1])
        + rmu[ni - 1]) * (t[ni] - ti));
        NumberOfBiopsy = k - 1;
        for (j = ni + 1; j <= NumberOfBiopsy; ++j)
        {
            r *= Math.Exp(-(rlan[j - 1] * Math.Exp(sumbb[i - 1]
            1)) + rmu[j - 1]) * (t[j] - t[
            j - 1]));
        }
        abs[i - 1] += r;
    }
}

abss = abs[i - 1] * 1000.0;
if (abss - (int)(abss) >= 0.5f)
{
    abss = (int)(abss) + 1.0;
}
else
{
    abss = (int)(abss);
}
abss /= 10.0;

retval = abs[i - 1];
return retval;
}

public static void PrintArray(double[] o, string Name)
{
    Console.WriteLine("-----Contents of {0}", Name);
    foreach (double d in o)
    {
        Console.WriteLine(string.Format("{0}", d.ToString("F5")));
    }
}

public static void PrintArray2(double[,] o, string Name)
{
    Console.WriteLine("-----Contents of {0}", Name);
    for (int i = 0; i <= o.GetUpperBound(0); ++i)
    {
        for (int j = 0; j <= o.GetUpperBound(1); ++j)
            Console.Write(string.Format("{0} ", o[i,
j].ToString("F2")));
        Console.WriteLine();
    }
}

#region Public Methods
#endregion
}
}

```

## APPENDIX B

### CONVERTING CLASS

```
using System;
using System.Text;
using System.Xml;

namespace Assessment_Tool_2.Engine
{
    public class BcptConvert
    {
        private const string _UNKNOWN = "UNKNOWN";
        private const string _UDERSCORE = "_";
        private const string _EMPTY = "";
        private const string _YES = "YES";
        private const string _NO = "NO";
        private const string _NA = "NA";
        private const string _ØBIRTHS = "Ø BIRTHS";

        /// Returns current age
        public static int GetCurrentAge(object o)
        {
            int rval;
            switch (o.ToString().ToUpper())
            {
                case _UNKNOWN:
                case _UDERSCORE:
                case _EMPTY:
                case _NA:
                    rval = 90;
                    break;
                case "< 35":
                case "<35":
                    rval = 34;
                    break;
                default:
                    rval = Convert.ToInt32(o);
                    break;
            }
            return rval;
        }

        /// Returns Projection Age
        public static int GetProjectionAge(string o)
        {
            int rval;
            switch (o.Trim().ToUpper())
            {
                case _UNKNOWN:
                case _UDERSCORE:
                case _EMPTY:
                case _NA:
                    rval = 90;
                    break;
                default:
                    rval = Convert.ToInt32(o);
                    break;
            }
            return rval;
        }

        /// Returns Menarche Age
        public static int GetMenarcheAge(object o)
        {
            int rval;
            switch (o.ToString().ToUpper())
            {
```

```

        case _UNKNOWN:
        case _UDERSCORE:
        case _EMPTY:
        case _NA:
        case _ØBIRTHS:
            rval = 99;
            break;
        case "7 TO 11":
            rval = 10;
            break;
        case "12 TO 13":
            rval = 13;
            break;
        case "> 13":
            rval = 15;
            break;
        default:
            rval = Convert.ToInt32(o);
            break;
    }

    return rval;
}

/// Returns First Live BirthAge
public static int GetFirstLiveBirthAge(object o)
{
    int rval;
    switch (o.ToString().ToUpper())
    {
        case _UNKNOWN:
        case _UDERSCORE:
        case _EMPTY:
        case _NA:
            rval = 99;
            break;

        case "NO BIRTHS":
            rval = 0;
            break;
        case "< 20":
            rval = 15;
            break;
        case "20 TO 24":
            rval = 22;
            break;
        case "25 TO 30":
            rval = 27;
            break;
        case "> 30":
            rval = 31;
            break;
        default:
            rval = Convert.ToInt32(o);
            break;
    }
    return rval;
}

/// Returns Number First Degree Relatives
public static int GetFirstDegRelatives(object o)
{
    int rval;
    switch (o.ToString().ToUpper())
    {

```

```

        case _UNKNOWN:
        case _UDERSCORE:
        case _EMPTY:
        case _NA:
            rval = 99;
            break;
        case "0":
            rval = 0;
            break;
        case "1":
            rval = 1;
            break;
        case "> 1":
            rval = 2;
            break;
        default:
            rval = Convert.ToInt32(o);
            break;
    }
    return rval;
}

/// Returns whether the person Ever Had Biopsy
public static int GetEverHadBiopsy(object o)
{
    int rval = 99;
    switch (o.ToString().ToUpper())
    {
        case _UNKNOWN:
        case _UDERSCORE:
        case _EMPTY:
        case _NA:
            rval = 99;
            break;
        case _NO:
        case "0":

        case "0":
            rval = 0;
            break;
        case _YES:
        case "1":
            rval = 1;
            break;
    }
    return rval;
}

/// Returns Number of Biopsy
public static int GetNumberOfBiopsy(object o)
{
    int rval;
    switch (o.ToString().ToUpper())
    {
        case _UNKNOWN:
        case _UDERSCORE:
        case _EMPTY:
        case _NA:
            rval = 99;
            break;
        case "1":
            rval = 1;
            break;
        case "> 1":
            rval = 2;
            break;

        default:
            rval = Convert.ToInt32(o);
            break;
    }
    return rval;
}

```

```

/// Returns whether the person had Hyper Plasia or not
public static int GetHyperPlasia(object o)
{
    int rval;
    switch (o.ToString().ToUpper())
    {
        case _UNKNOWN:
        case _UDERSCORE:
        case _EMPTY:
        case _NA:
            rval = 99;
            break;
        case _NO:
            rval = 0;
            break;
        case _YES:
            rval = 1;
            break;
        default:
            rval = Convert.ToInt32(o);
            break;
    }
    return rval;
}

/// Returns Current Age Indicator/Index
public static int CurrentAgeIndicator(int currentAge)
{
    int rval = 0;
    if (currentAge < 50)
        rval = 0;
    else if (currentAge >= 50)
        rval = 1;
    return rval;
}

/// gets the menarche age
public static int MenarcheAge(int menarcheAge)
{
    int rval = 0;
    if (menarcheAge >= 7 && menarcheAge < 12)
        rval = 2;
    else if (menarcheAge >= 12 && menarcheAge < 14)
        rval = 1;
    else if (menarcheAge >= 14 && menarcheAge <= 39 || menarcheAge == 99)
        rval = 0;
    return rval;
}

/// Returns First Live Birth Age
public static int FirstLiveBirthAge(int firstLiveBirthAge)
{
    int rval = 0;
    if (firstLiveBirthAge == 0)
    {
        // no live birth
        rval = 2;
    }
    else if (firstLiveBirthAge > 0)
    {
        if (firstLiveBirthAge < 20 || firstLiveBirthAge == 99)
            rval = 0;
        else if (firstLiveBirthAge >= 20 && firstLiveBirthAge < 25)
            rval = 1;
        else if (firstLiveBirthAge >= 25 && firstLiveBirthAge < 30)
            rval = 2;
        else if (firstLiveBirthAge >= 30 && firstLiveBirthAge <= 55)
            rval = 3;
    }
    return rval;
}

```

```

/// Returns Number of first degree relatives
public static int FirstDegRelatives(int firstDegRelatives)
{
    int rval = 0;
    if (firstDegRelatives == 0 || firstDegRelatives == 99)
        rval = 0;
    else if (firstDegRelatives == 1)
        rval = 1;
    else if (firstDegRelatives >= 2 && firstDegRelatives <= 31)
        rval = 2;
    return rval;
}

/// Returns Number of first degree relatives based on race
public static int FirstDegRelatives(int firstDegRelatives, int race)
{
    int rval = 0;
    if (firstDegRelatives == 0 || firstDegRelatives == 99)
        rval = 0;
    else if (firstDegRelatives == 1)
        rval = 1;
    else if (firstDegRelatives >= 2 && firstDegRelatives <= 31 && race < 7)
        rval = 2;
    else if (firstDegRelatives >= 2 && race >= 7)
        rval = 1;
    return rval;
}

/// Returns whether a woman ever had biopsy or not
public static int EverHadBiopsy(int everHadBiopsy)
{
    int rval = 0;
    switch (everHadBiopsy)
    {
        case 99:
            //case 0:
            rval = 0;
            break;
        default:
            rval = everHadBiopsy;
            break;
    }
    return rval;
}

/// Returns Number Of Biopsies
public static int NumberOfBiopsy(int numberOfPreviousBiopsy, int everHadBiopsy)
{
    int rval = 0;
    if (everHadBiopsy == 99)

        rval = 99;
    else if (numberOfPreviousBiopsy== 0 || (numberOfPreviousBiopsy == 99 && everHadBiopsy == 99))
        rval = 0;
    else if (numberOfPreviousBiopsy== 1 || (numberOfPreviousBiopsy == 99 && everHadBiopsy == 1))
        rval = 1;
    else if (numberOfPreviousBiopsy > 1 && numberOfPreviousBiopsy <= 30)
        rval = 2;
    return rval;
}

```

```

/// Returns Hyperlasia value
public static int Hyperplasia(int hyperplasia, int everHadBiopsy)
{
    int rval;
    if (everHadBiopsy == 0)
        rval = 99;
    else
    {
        switch (hyperplasia)
        {
            default:
                rval = hyperplasia;
                break;
        }
    }
    return rval;
}

/// Returns RHyperplasia value
public static double RHyperplasia(int hyperplasia, int everHadBiopsy)
{
    double rval;
    //if (everHadBiopsy == 0) hyperplasia = 0;
    switch (hyperplasia)
    {
        case 1:                // hyperplasia=yes
            rval = 1.82;
            break;
        case 0:                // hyperplasia=no
            rval = 0.93;
            break;
        default:               // hyperplasia=never had biopsy
            rval = 1.0;
            break;
    }
    return rval;
}
}
}
}

```

## APPENDIX C

### RISK CLASS

```
using System;
using System.Collections.Generic;
using System.Text;
using Assessment_Tool_2.Engine;

namespace Assessment_Tool_2
{
    internal static class Helper
    {
        internal static void RiskCalc(
            int riskIndex,
            int currentAge,
            int projectionAge,
            int menarcheAge,
            int firstLiveBirthAge,
            int everHadBiopsy,
            int numberOfBiopsy,
            int hyperPlasia,
            int firstDegRelatives,
            int race,
            out double absoluteRisk,
            out double averageRisk)
        {
            int ageIndicator;
            double rHyperPlasia;

            currentAge = BcptConvert.GetCurrentAge(currentAge);
            menarcheAge = BcptConvert.MenarcheAge(menarcheAge);
            firstLiveBirthAge = BcptConvert.FirstLiveBirthAge(firstLiveBirthAge);
            everHadBiopsy = BcptConvert.EverHadBiopsy(everHadBiopsy);
            numberOfBiopsy = BcptConvert.NumberOfBiopsy(numberOfBiopsy, everHadBiopsy);
            hyperPlasia = BcptConvert.Hyperplasia(hyperPlasia, everHadBiopsy);
            firstDegRelatives = BcptConvert.FirstDegRelatives(firstDegRelatives);
            race = BcptConvert.GetRace(race.ToString());

            if (race < 7)
                firstDegRelatives = BcptConvert.FirstDegRelatives(firstDegRelatives);
            else
                firstDegRelatives = BcptConvert.FirstDegRelatives(firstDegRelatives,
                    race);

            ageIndicator = BcptConvert.CurrentAgeIndicator(currentAge);
            rHyperPlasia = BcptConvert.RHyperplasia(hyperPlasia, everHadBiopsy);

            riskIndex = 1; //get absolute risk
            RiskCalculator oBcpt = new RiskCalculator();
            absoluteRisk = oBcpt.CalculateAbsoluteRisk(
                currentAge // [t1]
                , projectionAge // [t2]
                , ageIndicator // [i0]
                , numberOfBiopsy // [i2]
                , menarcheAge // [i1]
                , firstLiveBirthAge // [i3]
                , firstDegRelatives // [i4]
                , everHadBiopsy // [iever]
                , hyperPlasia // [ihyp]
                , rHyperPlasia // [rhyp]
                , race // [race]
            );
        }
    }
}
```



```

if (race < 7)
    firstDegRelatives = BcptConvert.FirstDegRelatives(firstDegRelatives);
else
    firstDegRelatives = BcptConvert.FirstDegRelatives(firstDegRelatives,
        race);

ageIndicator = BcptConvert.CurrentAgeIndicator(currentAge);
rHyperPlasia = BcptConvert.RHyperplasia(hyperPlasia, everHadBiopsy);

riskIndex = 1; //get absolute risk
RiskCalculator oBcpt = new RiskCalculator();
absoluteRisk = oBcpt.CalculateAbsoluteRisk(
    currentAge           //[t1]
    , projectionAge      //[t2]
    , ageIndicator        //[i0]
    , numberOfBiopsy     //[i2]
    , menarcheAge        //[i1]
    , firstLiveBirthAge  //[i3]
    , firstDegRelatives  //[i4]
    , everHadBiopsy      //[iever]
    , hyperPlasia        //[ihyp]
    , rHyperPlasia       //[rhyp]
    , race                //[race]
);

riskIndex = 2; //get average risk also
averageRisk = oBcpt.CalculateAverageRisk(
    currentAge           //[t1]
    , projectionAge      //[t2]
    , ageIndicator        //[i0]
    , numberOfBiopsy     //[i2]
    , menarcheAge        //[i1]
    , firstLiveBirthAge  //[i3]
    , firstDegRelatives  //[i4]
    , everHadBiopsy      //[iever]
    , hyperPlasia        //[ihyp]
    , rHyperPlasia       //[rhyp]
    , race                //[race]
);
}

internal static void CalcPercentage(double absoluteRisk, double averageRisk,
    out double absoluteRiskPctg, out double averageRiskPctg)
{
    absoluteRisk = Math.Round(absoluteRisk, 6);
    averageRisk = Math.Round(averageRisk, 6);
    absoluteRiskPctg = Math.Round(absoluteRisk * 100, 1);
    averageRiskPctg = Math.Round(averageRisk * 100, 1);
}
}
}

```

## APPENDIX D

### FORM 1 CODING

```
namespace Assessment_Tool_2
{
    public partial class Form1 : Form
    {
        public Form1()
        {
            InitializeComponent();
        }

        public void button1_Click(object sender, EventArgs e)
        {
            string path = @"D:\\demo.txt";

            dateTimePicker1.Text = DateTime.Now.ToShortDateString();

            string createText = "First Name : " + textBox1.Text
                + Environment.NewLine + "Last Name : " + textBox2.Text
                + Environment.NewLine + "Patient ID : " + textBox3.Text
                + Environment.NewLine + "Assessment Date : " +
                    dateTimePicker1.Text
                + Environment.NewLine + "-----"
                    + "-----";

            File.WriteAllText(path, createText);

            this.Hide();
            Form2 f2 = new Form2();
            f2.ShowDialog();
        }
    }
}
```

## APPENDIX E

### BCRA CLASS

```
namespace Assessment_Tool_2
{
    public partial class Form2 : Form
    {
        int age;
        int periodage;
        int birthage;
        int firstrelative;
        int undergobiop;
        int biopsy;
        int hyper;

        class userData
        {
            private int _value;
            private string _name;

            public int Value
            {
                get { return _value; }
                set { _value = value; }
            }

            public string Name
            {
                get { return _name; }
                set { _name = value; }
            }

            public userData(string name, int value)
            {
                _name = name;
                _value = value;
            }

            public override string ToString()
            {
                return _name;
            }
        }

        public Form2()
        {
            InitializeComponent();

            //currentAge
            comboBox1.Items.Add(new userData("35", 35));
            comboBox1.Items.Add(new userData("36", 36));
            comboBox1.Items.Add(new userData("37", 37));
            comboBox1.Items.Add(new userData("38", 38));
            comboBox1.Items.Add(new userData("39", 39));
            comboBox1.Items.Add(new userData("40", 40));
            comboBox1.Items.Add(new userData("41", 41));
            comboBox1.Items.Add(new userData("42", 42));
            comboBox1.Items.Add(new userData("43", 43));
            comboBox1.Items.Add(new userData("44", 44));
            comboBox1.Items.Add(new userData("45", 45));
            comboBox1.Items.Add(new userData("46", 46));
            comboBox1.Items.Add(new userData("47", 47));
            comboBox1.Items.Add(new userData("48", 48));
            comboBox1.Items.Add(new userData("49", 49));
            comboBox1.Items.Add(new userData("50", 50));
            comboBox1.Items.Add(new userData("51", 51));
        }
    }
}
```

```

comboBox1.Items.Add(new userData("52", 52));
comboBox1.Items.Add(new userData("53", 53));
comboBox1.Items.Add(new userData("54", 54));
comboBox1.Items.Add(new userData("55", 55));
comboBox1.Items.Add(new userData("56", 56));
comboBox1.Items.Add(new userData("57", 57));
comboBox1.Items.Add(new userData("48", 58));
comboBox1.Items.Add(new userData("59", 59));
comboBox1.Items.Add(new userData("60", 60));
comboBox1.Items.Add(new userData("61", 61));
comboBox1.Items.Add(new userData("62", 62));
comboBox1.Items.Add(new userData("63", 63));
comboBox1.Items.Add(new userData("64", 64));
comboBox1.Items.Add(new userData("65", 65));
comboBox1.Items.Add(new userData("66", 66));
comboBox1.Items.Add(new userData("67", 67));
comboBox1.Items.Add(new userData("68", 68));
comboBox1.Items.Add(new userData("69", 69));
comboBox1.Items.Add(new userData("70", 70));
comboBox1.Items.Add(new userData("71", 71));
comboBox1.Items.Add(new userData("72", 72));
comboBox1.Items.Add(new userData("73", 73));
comboBox1.Items.Add(new userData("74", 74));
comboBox1.Items.Add(new userData("75", 75));
comboBox1.Items.Add(new userData("76", 76));
comboBox1.Items.Add(new userData("77", 77));
comboBox1.Items.Add(new userData("78", 78));
comboBox1.Items.Add(new userData("79", 79));
comboBox1.Items.Add(new userData("80", 80));
comboBox1.Items.Add(new userData("81", 81));
comboBox1.Items.Add(new userData("82", 82));
comboBox1.Items.Add(new userData("83", 83));
comboBox1.Items.Add(new userData("84", 84));
comboBox1.Items.Add(new userData("85", 85));
comboBox1.Items.Add(new userData("86", 86));
comboBox1.Items.Add(new userData("87", 87));
comboBox1.Items.Add(new userData("88", 88));
comboBox1.Items.Add(new userData("89", 89));
comboBox1.Items.Add(new userData("90", 90));

//menarcheAge
comboBox2.Items.Add(new userData("Unknown", 99));
comboBox2.Items.Add(new userData("7", 7));
comboBox2.Items.Add(new userData("8", 8));
comboBox2.Items.Add(new userData("9", 9));
comboBox2.Items.Add(new userData("10", 10));
comboBox2.Items.Add(new userData("11", 11));
comboBox2.Items.Add(new userData("12", 12));
comboBox2.Items.Add(new userData("13", 13));
comboBox2.Items.Add(new userData("14", 14));
comboBox2.Items.Add(new userData("15", 15));
comboBox2.Items.Add(new userData("16", 16));
comboBox2.Items.Add(new userData("17", 17));
//firstLiveBirthAge
comboBox3.Items.Add(new userData("Unknown", 99));
comboBox3.Items.Add(new userData("No Birth", 26));
comboBox3.Items.Add(new userData("18", 18));
comboBox3.Items.Add(new userData("19", 19));
comboBox3.Items.Add(new userData("20", 20));
comboBox3.Items.Add(new userData("21", 21));
comboBox3.Items.Add(new userData("22", 22));
comboBox3.Items.Add(new userData("23", 23));
comboBox3.Items.Add(new userData("24", 24));
comboBox3.Items.Add(new userData("25", 25));
comboBox3.Items.Add(new userData("26", 26));
comboBox3.Items.Add(new userData("27", 27));
comboBox3.Items.Add(new userData("28", 28));
comboBox3.Items.Add(new userData("29", 29));
comboBox3.Items.Add(new userData("30", 30));
comboBox3.Items.Add(new userData("31", 31));
comboBox3.Items.Add(new userData("32", 32));

```

```

comboBox3.Items.Add(new userData("32", 32));
comboBox3.Items.Add(new userData("33", 33));
comboBox3.Items.Add(new userData("34", 34));
comboBox3.Items.Add(new userData("35", 35));
comboBox3.Items.Add(new userData("36", 36));
comboBox3.Items.Add(new userData("37", 37));
comboBox3.Items.Add(new userData("38", 38));
comboBox3.Items.Add(new userData("39", 39));
comboBox3.Items.Add(new userData("40", 40));
comboBox3.Items.Add(new userData("41", 41));
comboBox3.Items.Add(new userData("42", 42));
comboBox3.Items.Add(new userData("43", 43));

// firstDegreeRel
comboBox4.Items.Add(new userData("Unknown", 99));
comboBox4.Items.Add(new userData("0", 0));
comboBox4.Items.Add(new userData("1", 1));
comboBox4.Items.Add(new userData("2", 2));
comboBox4.Items.Add(new userData("3", 3));

//hadBiopsy
comboBox5.Items.Add(new userData("Unknown", 99));
comboBox5.Items.Add(new userData("No", 0));
comboBox5.Items.Add(new userData("Yes", 1));

//numBiopsy
comboBox6.Items.Add(new userData("Unknown", 99));
comboBox6.Items.Add(new userData("0", 0));
comboBox6.Items.Add(new userData("1", 1));
comboBox6.Items.Add(new userData("2", 2));

//hyperPlasia
comboBox7.Items.Add(new userData("Unknown", 99));
comboBox7.Items.Add(new userData("No", 0));
comboBox7.Items.Add(new userData("Yes", 1));
}
public void comboBox1_SelectedIndexChanged(object sender, EventArgs e)
{
    userData selectedData1 = (userData)comboBox1.SelectedItem;
    age = selectedData1.Value;
}

public void comboBox2_SelectedIndexChanged(object sender, EventArgs e)
{
    userData selectedData2 = (userData)comboBox2.SelectedItem;
    periodage = selectedData2.Value;
}

public void comboBox3_SelectedIndexChanged(object sender, EventArgs e)
{
    userData selectedData3 = (userData)comboBox3.SelectedItem;
    birthage = selectedData3.Value;
}

public void comboBox4_SelectedIndexChanged(object sender, EventArgs e)
{
    userData selectedData4 = (userData)comboBox4.SelectedItem;
    firstrelative = selectedData4.Value;
}

public void comboBox5_SelectedIndexChanged(object sender, EventArgs e)
{
    userData selectedData5 = (userData)comboBox5.SelectedItem;
    undergobiop = selectedData5.Value;
}

public void comboBox6_SelectedIndexChanged(object sender, EventArgs e)
{
    userData selectedData6 = (userData)comboBox6.SelectedItem;
    biopsy = selectedData6.Value;
}
}

```

```

public void comboBox7_SelectedIndexChanged(object sender, EventArgs e)
{
    userData selectedData7 = (userData)comboBox6.SelectedItem;
    hyper = selectedData7.Value;
}
public void button1_Click(object sender, EventArgs e)
{
    int currentAge = BcptConvert.GetCurrentAge(age);
    int menarchoAge = BcptConvert.GetMenarchoAge(periodage);
    int firstLiveBirthAge = BcptConvert.GetFirstLiveBirthAge(birthage);
    int firstDegreeRel = BcptConvert.GetFirstDegRelatives(firstrelative);
    int hadBiopsy = BcptConvert.GetEverHadBiopsy(undergobiop);
    int numBiopsy = BcptConvert.GetNumberOfBiopsy(biopsy);
    int hyperPlasia = BcptConvert.GetHyperPlasia(hyper);
    int race = BcptConvert.GetRace("1");

    double absRisk = 0, avgRisk = 0, absRiskPctg = 0, avgRiskPctg = 0;

    // Calculate 5 year risk.
    Helper.RiskCalc(0, currentAge, currentAge + 5, menarchoAge,
        firstLiveBirthAge, hadBiopsy, numBiopsy, hyperPlasia,
        firstDegreeRel, race, out absRisk, out avgRisk);

    Helper.CalcPercentage(absRisk, avgRisk, out absRiskPctg, out avgRiskPctg);

    // Calculate lifetime risk.
    Helper.RiskCalc(0, currentAge, 90, menarchoAge, firstLiveBirthAge, hadBiopsy,
        numBiopsy, hyperPlasia, firstDegreeRel, race, out absRisk,
        out avgRisk);
    Helper.CalcPercentage(absRisk, avgRisk, out absRiskPctg, out avgRiskPctg);

    MessageBox.Show("5 year risk" + Environment.NewLine + string.Format("This
        woman (age {0:N}) = {1:F}", currentAge, absRiskPctg)
        + Environment.NewLine + string.Format("Average woman (age
        {0:N}) = {1:F}", currentAge, avgRiskPctg)
        + Environment.NewLine
        + Environment.NewLine
        + Environment.NewLine
        + Environment.NewLine
        + "Lifetime risk" + Environment.NewLine + string.Format("This
        woman (to age 90): " + absRiskPctg.ToString("F1"))
        + Environment.NewLine + string.Format("Average woman (to age
        90): " + avgRiskPctg.ToString("F1")));

    string appendText = Environment.NewLine
        + Environment.NewLine
        + Environment.NewLine
        + Environment.NewLine
        + Environment.NewLine + "This is the result of the
        assessment"
        + Environment.NewLine
        + Environment.NewLine + "For 5 year risk assessment:"
        + Environment.NewLine + string.Format("At (age {0:N}),
        risk is {1:F} ", currentAge, absRiskPctg)
        + Environment.NewLine + string.Format("For average woman
        at (age {0:N}), risk is {1:F} ", currentAge,
        avgRiskPctg)
        + Environment.NewLine + "-----"
        + Environment.NewLine
        + Environment.NewLine
        + Environment.NewLine + "For lifetime risk assessment:"
        + Environment.NewLine + string.Format("Towards age 90,
        risk is " + absRiskPctg.ToString("F1"))
        + Environment.NewLine + string.Format("For average woman
        towards age 90, risk is " +
        avgRiskPctg.ToString("F1"));

    string path = @"D:\\demo.txt";
    File.AppendAllText(path, appendText);
}

```