North Carolina Agricultural and Technical State University
Aggie Digital Collections and Scholarship

Theses | Electronic Theses and Dissertations

Spring 2015

Predicting Hereditary Breast Cancer By Combining Family History And Medical Information

Sherin John
North Carolina Agricultural and Technical State University

Follow this and additional works at: https://digital.library.ncat.edu/theses

Recommended Citation

John, Sherin, "Predicting Hereditary Breast Cancer By Combining Family History And Medical Information" (2015). *Theses*. 258.
https://digital.library.ncat.edu/theses/258

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at Aggie Digital Collections and Scholarship. It has been accepted for inclusion in Theses by an authorized administrator of Aggie Digital Collections and Scholarship. For more information, please contact iyanna@ncat.edu.

Predicting Hereditary Breast Cancer by Combining Family History and Medical Information

Sherin John

North Carolina A&T State University

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department: Computer Systems Technology

Major: Information Technology

Major Professor: Dr. Rajeev Agrawal

Greensboro, North Carolina

2015

The Graduate School
North Carolina Agricultural and Technical State University

This is to certify that the Master's Thesis of

Sherin John

has met the thesis requirements of
North Carolina Agricultural and Technical State University

Greensboro, North Carolina
2015

Approved by:

Dr. Rajeev Agrawal Major Professor	Dr. Sambit Bhattacharya Committee Member
Dr. Evelyn Sowells Committee Member	Dr. Qing-An Zeng Committee Member

Dr. Clay Gloster
Department Chair

Dr. Sanjiv Sarin
Dean, The Graduate School

© Copyright by

Sherin John

2015

Biographical Sketch

Sherin John earned her Bachelor of Science (BSc) degree in Physics from Mahatma Gandhi University, India in 2000. She went on to pursue Master of Science (MSc) degree in Electronics from Mahatma Gandhi University and completed the degree requirements in 2005. In 2005, she achieved proficiency certification in Artificial Intelligence and Programming from Indian Institute of Science Bangalore. In 2013, she entered the Master of Science in Information Technology program at A&T. She is a member in Golden Key International Honor Society. In 2014, she got certification in Six Sigma Green Belt from A&T. While pursuing her degree, Sherin attempted and completed IBM's Master the Mainframe part 1 and part 2 challenges and entered her name in the wall of fame for the year 2014. Sherin has submitted her research at the IEEE SoutheastCon 2015 Conference and the paper has been accepted. She also presented her research paper at the 11th Annual Ronald E. McNair National Research Symposium.

Dedication

First and foremost, I dedicate this thesis to my Lord for His blessing that continue to flow into my life, and because of You, I made this through against all odds.

I would also like to thank my husband Freney John, daughter Aanjali Maria John and our parents for being very understanding, patient, and supportive during this journey. I also dedicate this thesis in the loving memory of my father John Joseph, who taught me invaluable lessons in life. It is their unconditional love that motivates me to set higher targets. I want to thank you all for the sacrifices of time and money you have made during my tenure in graduate school.

Acknowledgments

First, I would like to express my deepest gratitude and appreciation to my advisor Dr. Rajeev Agrawal, for his patience, motivation, enthusiasm, and immense knowledge. I am fortunate to have him as my mentor, and his guidance helped me in all the time of research and writing of this thesis. He always responded to my questions with a clear and timely response. His patience and support helped me overcome many crisis situations and finish the thesis. I would like to express my sincere gratitude to Dr. Clay Gloster, Jr. for being very supportive and providing me valuable guidance and reviews for my thesis works. I would also like to thank my other thesis committee members Dr. Sambit Bhattacharya (Department of Mathematics and Computer Science, Fayetteville State University), Dr. Evelyn Sowells, and Dr. Qing-An Zeng.

Table of Contents

List of Figures x

List of Tables xi

Abstract 1

CHAPTER 1 Introduction 2

1.1 What is Cancer? 5

1.2 What is Breast Cancer? 6

1.2.1 Factors that decrease the risk. 8

1.2.2 Factors that increase the risk. 8

1.3 Problem Statement 9

CHAPTER 2 Literature Review 11

2.1 Recent Breast Cancer Statistics 11

2.2 Open Source Cancer Risk Assessment Tools 12

2.2.1 Gail model – Breast cancer risk assessment tool. 13

2.2.2 IBIS tool. 13

2.2.3 Myriad II. 14

2.2.4 BOADICEA web application (BWA). 14

2.2.5 PENN II. 15

2.2.6 Breast cancer surveillance consortium risk calculator. 15

2.2.7 Hall detailed breast risk calculator. 15

2.2.8 Bright Pink. 16

2.3 Role of Family History in Breast Cancer Risk Assessment 16

CHAPTER 3 Different Models to Predict Breast Cancer 21

3.1 Gail Model 21

3.2 Claus Model 22
3.3 BRCAPRO Model 24
3.4 Cuzick-Tyrer Model 26
3.5 BOADICEA Model 27
3.6 Comparison of Different Models 28
CHAPTER 4 Analyzing Hereditary Information in Predicting Breast Cancer 32
4.1 Discussion on Gene Affecting Breast Cancer 33
4.1.1 BRCA1 and BRCA2. 35
4.1.2 PTEN. 37
4.1.3 Tp53. 39
4.2 Inheritance Patterns of BRCA1, BRCA2, Tp53 and PTEN Genes 41
4.2.1 Autosomal dominant inheritance. 42
4.2.2 Autosomal recessive inheritance. 42
4.2.3 X-linked dominant inheritance. 43
4.2.4 X-linked recessive inheritance. 43
4.3 How the Traits are Passing Down to the Generations? 44
CHAPTER 5 Prediction of Hereditary Breast Cancer 46
5.1 Flowchart to Predict the Hereditary Breast Cancer Risk 46
5.1.1 Checking BRCA1/BRCA2 mutation. 48
5.1.2 Checking Tp53 mutation. 54
5.1.3 Checking PTEN mutation. 56
5.2 Proposed Design 58
5.3 Software Tool to Predict Hereditary Breast Cancer in Current and Future Generation .. 61
5.3.1 Hereditary breast and ovarian cancer syndrome analysis. 61
5.3.1.1 Person with known mutation. 62

5.3.1.2 Personal history. 62

5.3.1.2.1 Personal history of breast cancer. 63

5.3.1.2.2 Personal history of epithelial ovarian cancer 64

5.3.1.2.3 Personal history of male breast cancer. 65

5.3.1.2.4 Personal history of pancreatic or prostate cancer. 65

5.3.1.3 Family history. 69

5.3.1.3.1 First- or second-degree relatives. 69

5.3.1.3.2 Third-degree relatives. 69

5.3.2 Li-Fraumeni syndrome analysis. 73

5.3.2.1 Individual from a family with known mutation. 73

5.3.2.2 Checking classic Li-Fraumeni Syndrome criteria. 74

5.3.2.3 Checking Chompret criteria. 74

5.3.2.4 Early-age-onset criteria. 75

5.3.3 Cowden syndrome analysis. 78

5.3.3.1 Individual with an already known mutation. 82

5.3.3.2 Personal history. 82

5.3.3.3 Family history. 82

5.3.4 Forecasting cancer in future offspring. 82

5.3.5 Output. 83

5.3.6 Features of the program. 85

CHAPTER 6 Future Work and Conclusion 87

References 88

List of Figures

Figure 1. Family health portrait. 17

Figure 2. Cell, chromosomes and DNA. 33

Figure 3. Autosomal dominant inheritance pattern. 44

Figure 4. BRCA1/BRCA2 personal history analysis. 48

Figure 5. BRCA1/BRCA2 personal history analysis (cont.). 49

Figure 6. BRCA1/BRCA2 family history analysis. 50

Figure 7. BRCA1/BRCA2 family history analysis (cont.). 51

Figure 8. Tp53 mutation analysis. 54

Figure 9. PTEN analysis. 56

Figure 10. PTEN analysis (cont.). 57

Figure 11. Proposed design. 59

Figure 12. A pedigree chart. 60

Figure 13. Image of the software program. 61

Figure 14. Personal history-BRCA1/BRCA2 analysis. 63

Figure 15. Family history-BRCA1/BRCA analysis. 69

Figure 16. Tp53 analysis. 75

Figure 17. PTEN analysis. 79

Figure 18. Output – BRCA1/BRCA2 mutation. 84

Figure 19. Output – PTEN mutation. 84

Figure 20. Output – Tp53 mutation. 85

Figure 21. Output – ‘not at risk’ case. 85

List of Tables

Table 1 Breast Cancer Statistics 11

Table 2 Comparison of Different Models 29

Table 3 Test Cases for Personal History-BRCA1/BRCA2 Analysis 66

Table 4 Test Cases for Family History –BRCA1/BRCA2 Analysis 70

Table 5 Test Cases for Tp53 Analysis 76

Table 6 Test Cases of PTEN Analysis 80

Abstract

Breast cancer is a malignant tumor, commonly found in women but rarely in men. One in eight US women will have the chance of developing breast cancer in their lifetime. A family history analysis and genetic test can identify the potential carrier gene that cause hereditary breast cancer. Mutated genes such as BRCA1, BRCA2, Tp53 and PTEN confer high lifetime risk of breast cancers and are the most significant genes for clinical studies. The study goals of this thesis were to: 1) develop an open source software tool to predict the risk of hereditary breast cancer and mutation risk of breast cancer in any person, 2) predict the risk of hereditary breast cancer in the next generation by analyzing the inheritance patterns of the genes, and 3) exploit medical guidelines and connect it with information technology.

In the program, we combine National Comprehensive Cancer Network (NCCN) guidelines of BRCA1/BRCA2, TP53, and PTEN gene mutations. We have sought the help of medical practitioners, medical literatures, and all other credential reports to validate the output of this research. The input data for this program are family history and other cancer risk assessment guidelines for the individual. We have developed this software tool in C++ and JavaScript. In the program, we identify a person's family history, age, and gender through questionnaires and make an evaluation of different possibilities and outcome. We analyze the risk of mutation with the help of family history. We have considered first, second, and third-degree relatives and individuals who belong to Ashkenazi Jewish ancestry.

After the analysis, and if the person is at risk, we predict the mutation risk of the next generation based on the gene's inheritance pattern. This tool would help the members of the general public who have the least knowledge about breast cancer and genetic inheritance, and it would also help them to identify the presence of mutated genes.

CHAPTER 1

Introduction

Medical data that is stored in computers present many promising advantages in medical research. Nowadays, medical reports and health information are becoming more available as the generations pass. Therefore, research that exploits medical data brings new advancements in medical systems. In this computer-driven world, genome mapping and medical technology have made it possible to get genetic information at low cost, which was not possible earlier. Advances in computer processing and medical technology made it possible to get a genome mapping of any person quickly. We can exploit these advancements to assess a person's risk factors for any given diseases that are inherited or related to a genetic defect. The Human Genome Project (HGP) is an international scientific research project with a primary goal of determining the sequence of chemical base pairs, which make up human DNA (Williams, 1997). It took efforts of several international institutes to complete, and remains the world's largest collaborative biological project. The HGP has identified about 20,500 human genes and provides information about the structure and whole functions of human genes. As the years go by, family health history stored in computers provides better assessments, accuracy, and standardization in forecasting diseases. Combining all known diseases, symptoms, and diagnoses helps medical practitioners to determine and forecast diseases in the current or next generations. Computers could be trained to forecast and predict specific diseases.

A family history gives a report about a person and his or her relatives (Agrawal, Suleiman, Seay, & Gloster, 2013). Families have many common factors, such as genes, lifestyle, environment and ethnic background. These common factors can be related to some medical conditions that run in a family. We can use genetic reports, family medical history, medical

reports, and lab reports to forecast a person's probability of getting diseases that are inherited. Getting genome mapping helps to diagnose some hereditary diseases. Even though, we cannot change the existing gene pattern, we can be informed and take precautions based on the available reports. Integrating medical, family, and genetic reports will help to predict hereditary diseases such as cancer, diabetics, and asthma. Furthermore, study and research about cancer is on the rise for various reasons. The World Health Organization estimates significant growth in cancer patients worldwide, and cancer is poised to overtake heart disease as the world's top killer. There are about two hundred cancers known at this time. The primary factor in cancer is the uncontrollable growth of cells. Cancer screening can be used to detect early symptoms of cancer. There are different reasons for cancer, and the primary reason is environmental factors, which include tobacco use, diet, obesity, infections, radiations, stress, lack of physical activity, and pollutants. Everyone has the risk of developing cancer in his or her lifetime. The factors behind developing cancer include age, lifestyle, chemicals, radiation, genetics, immune system, and infection. Most of the cancers are non-hereditary, and inherited genetic defect commonly causes hereditary cancers. Hereditary cancer is associated with highly penetrant genes, and pass to the next generation through different inheritance patterns. Hereditary cancer can be predicted earlier based on a pedigree analysis. A pedigree is a chart or diagram of family history analysis that includes parents, grandparents, and previous generations, and it checks the occurrence of any inherited diseases. The information collected from pedigree analysis is useful to determine the cancer risk for an individual or his/her progeny.

Cancer can be predicted in current or future generations based on the information gathered from family history tools, genetic and medical guidelines. Family health history plays an important role in determining the risk factors as it gives a link to hereditary, environmental,

cultural, and behavioral factors of the entire family (Aiello-Laws, 2011). The risk can be considered high, moderate, and average, based on the number of the family affected and age at the onset of the people who are affected in the family (Yoon et al., 2002). In this thesis, we have focused on predicting the risk of mutated genes that cause breast cancer, as most of the hereditary cancers are predictable at early stages.

Breast cancer can occur in both men and women, but it is far more common for women. A family history analysis and genetic test can identify the potential carrier gene behind breast cancer. Breast cancer is a malignant tumor which is capable of spreading into surrounding tissues or to distant areas. The risk factors involved with breast cancers are age, gender, genetic factors, family history, race/ethnicity, and habits. Breast cancer occurs in more often in females than in males because of the female hormones like estrogen and progesterone, which help cancer cell growth. The risk of developing breast cancer increases with age. Mutated genes such as BRCA1, BRCA2, PTEN, and Tp53 make the individual at high lifetime risk of breast cancer (Cancer Research UK, 2014). There are other genes such as PALB1, BRIP and LKB1 behind breast cancer, which contribute moderate risk. However, BRCA1, BRCA2, TP53, and PTEN are the genes considered most for clinical studies (Allain, 2008). The rest of the paper is organized as follows: Section 1.2 discusses the problem statement, and motivation for this thesis and Chapter 2 presents literature review. Chapter 3 explains different breast cancer prediction models available to predict cancer. Chapter 4 provides details about the analysis of hereditary information in predicting breast cancer. Chapter 5 explains the prediction of hereditary breast cancer risk in current and next generations, and Chapter 6 discusses future work.

1.1 What is Cancer?

Cancer is a general term used to define abnormal cell division, and it attacks other tissues. Cancer spreads into the body through the blood and lymphatic systems. The cell is the basic unit of our bodies. All living things are made up of cells. The human body is made up of different kinds of cells. Different cells have different functions in the body (Williams, 1997). Like other living things, cells die, and new cells will be produced by the body. Normally cells divide faster in early ages and allow the body to grow. After a person reaches adulthood, cell division happens only to replace the dying cells.

DNA is the genetic material of a cell, and it controls growth and cell division. Sometimes when the DNA becomes damaged, cell division and growth go out of control, and that can lead to accumulations of extra cells called tumors. Abnormal cells or cancer cells can spread to other body parts through the blood or lymphatic systems. Spreading cancer into other parts of the body is called metastasis (National Cancer Institute [NCI], 2015). Not all tumors are cancerous. Tumors can be divided into two types, benign tumors, and malignant tumors. Benign tumors are not cancerous and do not invade other parts of the body. Malignant tumors are cancerous, and they are able to attack other parts of the body. Cancers can be divided into different categories, and the main categories are as follows:

- Carcinoma: Cancer that begins in the skin or in the tissues that cover the other organs in the body is called carcinoma cancer.
- Sarcoma: Cancer that originates from bone, fat, muscle, blood vessels, cartilage, or other connective tissues is called sarcoma.
- Leukemia: Cancer that begins in the bone marrow is known as leukemia. Bone marrow is a blood-forming tissue, found in the center of the bones, and it produces blood cells.

- Lymphoma and Myeloma: These are cancers that affect the immune system.
- Central nervous system cancers: These are cancers that are related to the brain and spinal cord (NCI, 2015).

The cancer risk assessment can be used to determine the level of risk for each individual. Many family history factors can give details about cancer risk. Identifying cancer with genetic testing can be used for treatment and for other prevention methods for patients. Therefore, cancer risk assessment and testing can be useful in personalizing therapies to improve the outcomes of the treatments (Riley et al., 2012). Cancer risk assessment is composed of many factors, including personal and family health history, reproductive history and hormone use, environmental risk factors, lifestyle and genetic information (Aiello-Laws, 2011). An accurate risk assessment directs the proper use of cancer screening, risk reduction, genetic testing and early detection strategies. It is also important to asses a complete medical and surgical history, including current medications, lifestyle, environmental exposures, and congenital anomalies. If the patient is a female, a detailed reproductive history is needed (Aiello-Laws, 2011). In order to get a detailed assessment, we can use pedigree analysis, genetic testing, risk modeling, biochemical testing and imaging, physical features, potential hereditary syndromes and cancer risk assessment for a patient's biological relatives. The information gathered can be used for cancer screening, prevention, and risk reduction, and also for identifying cancer at-risk family members (Riley at al., 2012).

1.2 What is Breast Cancer?

Breast cancer is a cancer that is found in breast cells. It is commonly found in women and rarely in men. Female breasts are made up of milk-producing glands, lobules, and ducts that carry milk from lobules to nipples. Ducts and lobules are covered with fatty connective tissue

called stroma. Breast cancer that begins in the tissues of ducts is called ductal cancers, and cancer that begins in the tissues of lobules is known as lobular cancer (American Cancer Society, 2014). Breast cancer can spread to other body parts through the lymphatic system. Lymph cells are immune system cell that helps the cells to fight infections and foreign materials that enter the body. A lymphatic system consists of lymph cells, lymph fluid, and lymph vessels that carry the lymph around the body. Breast cancer cells are able to invade the lymph vessels and begin to grow in the lymph nodes. The lymph vessels carry these cancerous cells to the other part of the body, which is known as metastatic or secondary cancer (American Cancer Society, 2014).

The risk of developing breast cancer in women increases with their ages. The other factors that increase the risk are hormone replacement factors, which include the use of hormone replacement therapy (HRT), hormonal contraceptives, obesity, and alcohol. The risk is reduced by different factors, which are being a young age at first childbirth, breast feeding, late menarche and early menopause. The risk will double if the women have an affected first-degree relative compared to women who don't have a family history of breast cancer. High penetrance genes such as BRCA1 and BRCA2 contribute high risk but only account for a small percentage of breast cancer. The lifetime risk of breast cancer with BRCA1 accounts for 85%, and ovarian cancer accounts 40% (Murray & Davies, 2013). However, the BRCA2 pathogenic variant shows a little lower risk, especially for ovarian cancer, and it accounts less than 20%. For males BRCA mutations contribute a risk for prostate cancer and breast cancer that is only 6 % (Murray & Davies, 2013). Genetic information or family history can be identified using a family tree diagram or pedigree analysis. Patients are assigned to one of the risk categories; women at or near population risk of developing breast cancer, women at raised risk of developing breast cancer and women at high risk of developing breast cancer. Many factors can be found in family

histories that are suggestive of BRCA pathogenic variant. The clues include, multiple individuals on one side of the family with breast or ovarian cancer, a young average age of diagnosis, male breast cancer, associated cancer, and individuals with more than one primary cancer, including bilateral breast cancer, and Ashkenazi Jewish Ancestry (Murray & Davies, 2013).

1.2.1 Factors that decrease the risk. The main risk factors that affect breast cancer include age, gender, and mutation in breast cancer genes. A woman who started her first menstrual period before the age of 12 is slightly at risk for breast cancer. The three hormones that affect any woman in monthly menstrual cycles are estrogen, testosterone, and progesterone. Estrogen level will be high in the body during the week of the menstrual cycle. Estrogen level will be low after the menstrual days. Exposure of breast tissue to increased levels of estrogen makes the women more at risk of developing breast cancer. Consequently, starting menstrual period at an early age increases the risk of breast cancer (Centers for Disease Control and Prevention [CDC], 2014).

Before menopause, estrogen is produced by the ovaries, and after menopause, estrogen will be produced by fat tissue. Women who have menopause at an early age are less exposed to estrogen. Giving birth to more children, giving birth to your first child at a young age, and breastfeeding your children can also decrease the risk. Pregnancy and breast-feeding reduces the risk of breast cancer as it reduces the women's lifetime exposure to menstrual cycles and hence, reduces exposure to estrogen level. Physical activity lowers the level of estrogen in the blood and boosts the immune system of the body. Becoming obese after menopause may increase the risk of breast cancer, as it increases the risk of estrogen receptor-positive breast cancer (CDC, 2014).

1.2.2 Factors that increase the risk. Long-term use of hormone replacement therapy: Women use hormone replacement therapy to ease menopause. The hormones commonly

prescribed in the therapy are estrogen and progesterone or progestin. Studies show that a woman taking HRT has a high risk of developing breast cancer. If you have ever been diagnosed with breast cancer, the recurrence of cancer in another part of the breast is always high. If you have first-, second-, or third-degree relatives diagnosed with breast cancer, the risk of inheriting the mutated gene is always possible, and that makes you at risk. In rare cases, radiation can cause secondary breast cancer. Diethylstilbestrol is a form of estrogen hormone. This hormone was prescribed to women to avoid miscarriage, premature labor, and related pregnancy problems between 1940 and 1971. Different studies (Palmer et al., 2002) show that the women who have received DES are more at risk of breast cancer when compared with women who never had DES. Women who have high breast density are more at risk than women who have less breast density (CDC, 2014). Alcohol increases the level of estrogen and other hormones relate to hormone-receptor- positive breast cancer. During the night, our body produces the hormone melatonin that makes us feel sleepy and want to a rest. Exposure to artificial light, during the night, may reduce the level of melatonin. As a result, it causes the body to produce more estrogen and makes people at risk of breast cancer.

1.3 Problem Statement

According to Mayo Clinic, after skin cancer, breast cancer is the most common cancer diagnosed in women in the United States. We have chosen breast cancer mutation risk, as breast cancer is considered one of the leading causes of cancer deaths in women. As a common cancer in women, are we able to predict the risk of carrying mutated genes based on the family history? Family history can be considered one of the important risk factors for many chronic diseases. The cancer risk assessment includes a complete personal medical history and previous generations' medical history to the extent that they are available. Family history helps to identify

higher risk for some diseases and also helps to get an idea about early warning signs of diseases. Does family history play an important role in predicting hereditary cancers? There are different models available to the public to check breast cancer risk. Can we completely depend on these prediction models to analyze the mutation risk? Do these models analyze complete family history and other non-hereditary factors that affect sporadic breast cancer such as birth control, hormone therapy, breastfeeding, drinking alcohol, and physical activity? In hereditary breast cancer, the gene passes to the next generation through an autosomal dominant inheritance pattern. Every child inherits two copies from their parents; one from their mother and one from the father. As BRCA1, BRCA2, Tp53, and PTEN genes exhibit autosomal inheritance pattern; there are four possible combinations for the child's gene copy. By analyzing the inheritance patterns, how can we predict the risk of carrying mutated genes in the offspring?

The aim of this research is to develop a software tool to predict the risk of hereditary breast cancer in any living person and also in their children. The backbone of this software program is NCCN guidelines. The existing tools to check the hereditary breast cancer are only considering BRCA1/BRCA2 genes. Some of the tools are not considering family history. In this research, we have included PTEN and Tp53 genes and also considered first-, second-, and third-degree relatives. NCCN guidelines are the most thorough, comprehensive, up-to-date, and reliable recommendations and are from a group of worldwide clinical practitioners.

CHAPTER 2

Literature Review

2.1 Recent Breast Cancer Statistics

There are more than two hundred types of cancers that are recognized in the human body. Nowadays, study and research about cancer is on the rise for various reasons. The World Health Organization estimates significant growth in cancer patients worldwide and that cancer is poised to overtake heart disease as the world's top killer. According to the American Cancer Society, the most common type cancer is breast cancer, with about 235,000 new cases expected in the United States in 2014 (American Cancer Society, 2015). The following table shows the estimated numbers of new cases and death for each common cancer.

Table 1

Breast Cancer Statistics

Cancer Type	Estimated New Cases	Estimated Deaths
Bladder	74,690	15,580
Breast (Female – Male)	232,670 – 2,360	40,000 – 430
Colon and Rectal (Combined)	136,830	50,310
Endometrial	52,630	8,590
Kidney (Renal Cell and Renal Pelvis) Cancer	63,920	13,860
Leukemia (All Types)	52,380	24,090
Lung (Including Bronchus)	224,210	159,260
Melanoma	76,100	9,710
Non-Hodgkin Lymphoma	70,800	18,990

Table 1

Cont.

Pancreatic	46,420	39,590
Prostate	233,000	29,480
Thyroid	62,980	1,890

In the United States 12% (one in eight) of women are at risk of invasive breast cancer during their lifetime (Breastcancer.org, 2014). About 62,570 new cases of non-invasive breast cancer have been reported in the United States. Breast cancer incidence rate was decreased from the year 2002 to 2003 after the decrease in the use of hormone replacement therapy. The study shows a large decrease in the death rate after 1989, and it is thought to be the result of early detection and prevention. However, the death rate is higher than any other cancer except lung cancer. White women have more probability to develop breast cancer than African American women. However, breast cancer under the age of 45 is more prevalent in African American women than white women. Asian, Native American, and Hispanic women have a lower risk of developing breast cancer. In 2014, recent data shows 2.8 million women are identified with breast cancer, which includes women who are undergoing treatment and who finished treatment. Among women with breast cancer, 15% have a family history of breast cancer (American Cancer Society, 2015).

2.2 Open Source Cancer Risk Assessment Tools

There are many software programs available to assess breast cancer risk and screening. Early detection of breast cancer helps to make a decision about the prevention of the cancer. None of the tools are accurate. However, it gives an understanding of breast cancer risk based on

your habits and family history. These programs have many features and limitations. One of the main drawbacks of all the programs is, none of the tool takes all risk factors associated with breast cancer. Each tool calculates only an estimated result based on the risk factors it considers. As there is no comprehensive breast cancer risk assessment tool, it is advisable to consider other health care assist to detect breast cancer earlier. Before checking the risk, try to use a reliable and credible program that has been created by a faithful source. Select a tool that provide the features and limitations of the tool, and understand how it was developed.

2.2.1 Gail model – Breast cancer risk assessment tool. This program has been developed (http://www.cancer.gov/BCRISKTOOL/) by scientists at the National Cancer Institute (NCI) and it is periodically updated with new data. The Gail model is used to calculate invasive breast cancer risk over a five year period and a lifetime period (Gail et al., 1999). This tool is useful for healthcare providers to assess the breast cancer risk and also help to make decisions in prevention strategies. This tool mainly focuses on non-genetic factors, personal history, previous breast biopsies, age of menarche, age of first live birth and family history of first-degree relatives. The main drawbacks of the program are, it does not consider other breast cancer-related factors such as breast density, hormone replacement therapy, breastfeeding, smoking, age at menopause and alcohol consumption.

2.2.2 IBIS tool. The IBIS tool (http://www.ems-trials.org/riskevaluator/) is used to calculate a person's likelihood of carrying the BRCA 1 or 2 mutations, which are associated with increased breast cancer risk. It estimates the likelihood of a woman developing breast cancer in ten years and over the course of her lifetime. The tool is used to help inform a person's decision-making about genetic counseling and testing.

The IBIS tool is developed based on the Tyrer-Cuzick model and it predicts a person's risk of carrying BRCA1 or BRCA2 genes. It calculates the risk of developing breast cancer over ten years and a lifetime period based on the estimation of BRCA1/BRCA2 genes. It is a good breast cancer decision-making tool. It considers many risk factors such as age, age at first live birth, body mass index, hormone replacement therapy, age at menopause, and family history. This tool suggests genetic counseling if the risk is 10% or higher. The disadvantages of the tool are that it does not consider breast density, alcohol consumption, oral contraceptive pill use, etc.

2.2.3 Myriad II. Myriad II (http://www.myriadpro.com/brca-risk-calculator/calc.html) has been developed based on empirical data which was conducted among ten thousand women with germline mutations in BRCA1/BRCA2 genes. The Myriad II prevalence tables (Frank et al., 1998) are based on proband and family history (van et al., 2010). The tool considers different risk factors such as relatives with ovarian cancer at any age, relatives with breast cancer, age of the relatives, and male breast cancer. It allows only a maximum of three relatives and does not give separate result for BRCA1 and BRCA2. It gives a separate analysis for Ashkenazi Jewish Ancestry.

2.2.4 BOADICEA web application (BWA). The BOADICEA is another tool, and it has collaborated with MENDEL, (http://ccge.medschl.cam.ac.uk/boadicea/advice-for-the-public/) which is pedigree software used to estimate the risks of breast and ovarian cancer in women. It calculates the probability of BRCA1 or BRCA2 genes. It also predicts the possibility of a woman developing breast cancer in five years and in her lifetime. The tool uses family history and gene mutation risk to calculate the risk of breast cancer. The tool recommends genetic counselling if the risk is 10% or higher. This tool only considers mutation status and family history. It doesn't consider any other non-genetic factors. The tool was developed based on data from the United

Kingdom. Therefore, the calculation will have less accuracy for other people from other countries.

2.2.5 PENN II. It is a free web-based tool (http://www.afcri.upenn.edu/itacc/penn2/), and it uses logistic regression based on pedigrees from Europe and North America. The tool considers age at onset of breast cancer, ovarian cancer, bilateral breast cancer, pancreatic cancer, and prostate cancer. The results of the program contain both BRCA1 and BRCA2 mutations. The results don't contain breast cancer risk, and it recommends genetic testing for people with 5% or higher risk. It gives a one-page questionnaire, and the tool considers first-, second-, and third-degree relatives. It excludes proband from families with no breast cancer.

2.2.6 Breast cancer surveillance consortium risk calculator. The Breast Cancer Surveillance Consortium (BCSC) risk calculator was created by scientists who participate in Breast Cancer Surveillance Consortium (https://tools.bcsc-scc.org/BC5yearRisk/intro.htm). This model gives a calculation of women's risk over a five-year time period. The tool was designed for health care professionals. The tool does not consider previous diagnosis of breast cancer, lobular carcinoma in situ, ductal carcinoma in situ, atypical ductal hyperplasia, or breast augmentation. This tool only considers people who are between the ages of 35 and 79. The model focuses on age, race, ethnicity, family history of first-degree relatives, history of breast biopsy, and breast density.

2.2.7 Hall detailed breast risk calculator. This open source breast cancer risk calculator was designed by Dr. Halls (http://halls.md/breast/risk.htm). This risk assessment tool includes many factors that are not included in the Gail model. This tool underestimates various factors such as two or more people in the same generation with the same cancer, breast cancer before the age of 50, ovarian cancer, cancers in both breasts, and melanoma. The tool includes first-degree

relatives, breast biopsies, age at menarche, age at first live birth, current age, and race. The result of the estimation falls into different categories. It gives the risk over five years, ten years, twenty years, thirty years, and a lifetime period.

2.2.8 Bright Pink. Bright Pink is (http://www.brightpink.org/knowledge-is-power/assess-your-risk/) another open source tool that calculates breast cancer risk. It is designed for academic use. This tool considers first- and second-degree relative breast cancer history, age at onset of breast cancer, multiple breast cancers, male breast cancer, ovarian cancer, other related cancers, and known mutations. The tool also considers non-genetic factors such as body mass index, alcohol consumption, smoking, physical activity, diet, birth control pills, and age at first live birth. The result is given in a PDF format.

The above open source programs have many limitations. These tools are only considering BRCA1 and BRCA2 mutations. None of the existing breast cancer tools available considers other high-risk genes such as PTEN and Tp53. Many of the existing tools are proprietary and not available to the public. Most of these existing tools have been developed based on the data collected from a particular demography or ethnicity. Many of the tools available do not consider third-degree relatives and family history.

2.3 Role of Family History in Breast Cancer Risk Assessment

Family history can be considered one of the important risk factors for many chronic diseases. The cancer risk assessment includes a complete personal medical history and previous generations' medical history to the extent that they are available. We have developed a family history diagram, as shown in Figure 1, with the help of the US Department of Health and Human Service website (HHS, 2014).

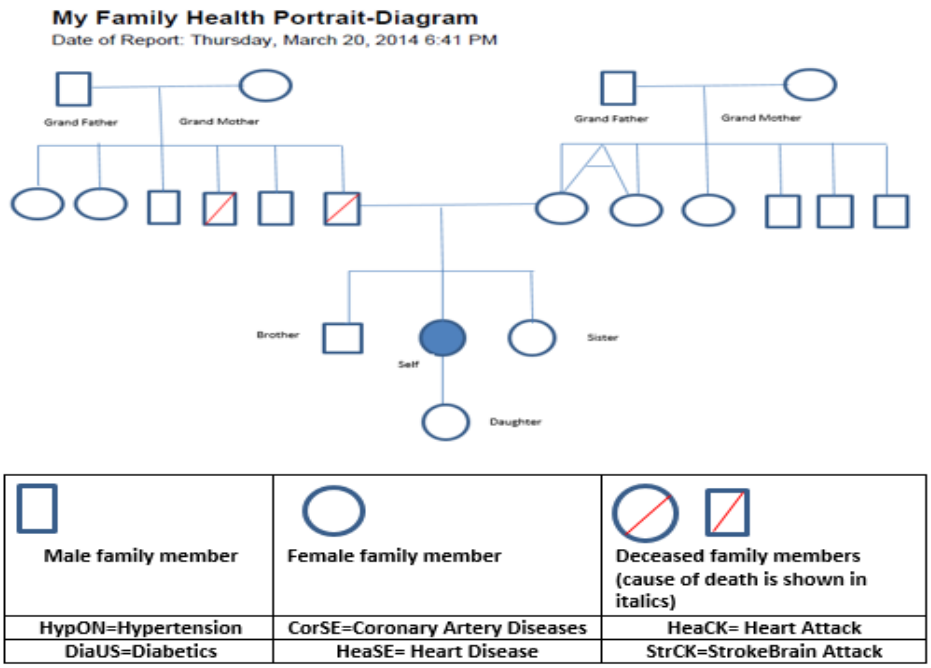

Figure 1. Family health portrait.

The diagram helps a person to store his/her entire family history as a report or a diagram. The diagram assembles all information to make a pedigree analysis and give a family tree based on health records. It helps to identify the higher risk for some diseases and also helps to get an idea about early warning signs of diseases (HHS, 2014). This health information can be stored in HealthVault, and it will be accessible to the person at any time. In this project, this family history diagram makes it easy to develop the program based on pedigree analysis.

Figure 1 describes the entire family history of a person that includes history of the deceased people and the reasons for death. It gives an overall idea about a person's family health history that includes parents, grandparents, and first-degree relatives. It will be helpful to analyze if a person has any risk of cancer. If this family history diagram contains any repeated diseases among first-degree or second-degree relatives, the chances of getting the disease in the next

generation are high. Many family history factors can give details about hereditary diseases such as cancer, diabetics, heart attack, and asthma.

Family history is an essential component of any cancer risk assessment. Family history tool must be simple, easily applied and inexpensive, and should be capable of identifying a person according with the intensity of the diseases and positively influencing healthy behaviors. When it comes to chronic diseases such as cancer and cardio vascular disease, family history reflects hereditary factors, environmental, cultural and behavioral factors (Yoon, Scheuner, & Khoury, 2003). A typical family history should be obtained from the proband. For each individual, specifically biological relationships among family members, ages, causes of death, genetics, surgical history, and any other medication should be marked. Hereditary cancer is associated with highly penetrant genes, and it reflects the range of risk associated with a mutation in the gene. Penetrance data from families with single gene mutation hereditary cancer syndromes are revised as modifier genes. Oncology professionals should reassess the family history, as family histories are dynamic. Accuracy of risk assessment can be affected by different factors (Aiello-Laws, 2011). The cancer risk assessment can be used to determine the risk level for each individual. Based on the family history, if an individual has a suggestive inherited cancer syndrome, a genetic testing can be carried out to detect the repeated cancer present in the family. The risk assessment includes a complete personal medical history and previous generations' medical history, if available. A standardized pedigree nomenclature and well-prepared questions are important factors in assessing the risk. Before starting questionnaires, preparing a family health history portrait of the individual will help in the risk assessment (Riley et al., 2012). Genetic tumor or cancer risk assessment reveals the possibility of DNA alterations, which are strong causes of cancer. Genetic risk assessment involves identifying and counseling

individuals at increased risk for cancer development and distinguishing the intensity of cancer. Genetic testing can be considered when an individual has a family history of genetic cancer susceptibility. Genetic testing is most useful when initiated when the person with cancer concerns is diagnosed at a young age. The test should begin by testing for the syndrome that the individual's features most resemble (Aiello-Laws, 2011). In order to get a detailed assessment, we can use pedigree analysis of the entire family. The information gathered could be used for cancer screening, prevention, and risk reduction, and also for identifying cancer at-risk family members (Lalloo & Evans, 2012).

Family health history is a significant cause that increases a person's risk of developing breast and ovarian cancer. Every individual should be mindful of these risk factors in his/her family. In general, the number of first-, second-, or third-degree relatives who have or had breast or ovarian cancer, and the age at onset of their diagnosis, increases the person's breast cancer risk. If you have a first-degree relative (sister, brother, father, son, or daughter) identified with breast cancer, the risk will be doubled (breastcancer.org, 2014). At the same time, if you have two first-degree relatives, the risk will be five times higher than the average. In many cases, family history of breast cancer is related to abnormal genes such as BRCA1, BRCA2, PTEN, and Tp53.

Different studies among families show that many women developed breast and/or ovarian cancer together. A majority of the cases are reported at earlier than usual ages. Medical practitioners suggested the term Hereditary Breast and Ovarian Cancer Syndrome (HBOC) after the analysis in many families. The mutation in genes BRCA1 and BRCA2 are the main causes of this disease. Even though mutations in either gene can cause the risk, the BRCA1 mutation causes a higher risk of breast cancer. The BRCA2 mutation is more likely related to male breast

cancer, pancreatic cancer, and prostate cancer. The BRCA gene mutation is common in people who belong to Ashkenazi Jewish families. There are other hereditary diseases such as Cowden syndrome (gene Tp53 mutation) and Li-Fraumeni syndrome (PTEN gene mutation) that have a high impact on breast cancer (American Cancer Society, 2014).

Women with a family history of breast/ovarian cancer can go through genetic counseling to calculate the risk of mutation in one of the BRCA genes. The professional will estimate the risk based on the personal or family history of cancer. After the estimation, if a mutation is present, the woman would have a high risk of breast/ovarian cancer. The calculation helps the person to find cancer early, and it will lower the risk of getting cancer. If the person is at risk of hereditary breast/ovarian cancer, it means that their close relatives have a 50% chance of having the same kind of mutation.

CHAPTER 3

Different Models to Predict Breast Cancer

Breast cancer risk assessment models help people in determining the risk. Medical practitioners use different models to assess the patient's risk of developing breast cancer. All the breast cancer tools that are available now have many drawbacks, and the accuracies of these models are not perfect. These models assess the risk either as the probability of developing breast cancer or the probability of mutated genes that cause breast cancer. It is important to assess the risk very accurately to make the right steps in prevention strategies, which range from lifestyle changes to breast removal. It is required to develop an accurate model to assess individual risk. The accuracy of the models depends on many factors such as lifestyle, diet, and hereditary history. However, breast cancer is likely to occur even in the absence of these known factors. This chapter discusses different breast cancer risk assessment models and also compares different models.

3.1 Gail Model

The Gail model was developed by Dr.Mitchell Gail, a senior investigator in a biostatistics department. The risk factors considered for the Gail model were age at menarche, age at first live birth, number of previous biopsies, and the number of first-degree relatives with breast cancer. The model is based on the relative risks of various combinations of these factors taken from the Breast Cancer Detection Demonstration Project (BCDDP). The cases used for this model consisted of women with breast cancer between 1978 and 1980 at twenty-nine participating centers that involved 280,000 women aged 35 to 74 years. This model groups the patients into two categories based on their age. The relative risks connected with previous breast biopsies were smaller for women of age 50 and older than women younger than 50. The model, again, is

subdivided into three parts: (a) Estimating the relative risk for a subject with a given group of risk factors at a given age compared to a subject without the identified risk factors; (b) estimating the baseline age-specific breast cancer hazard rate, which is the rate for a woman without identified risk factors; and (c) projecting the long-term probability of developing breast cancer on the basis of a consideration of competing risks and the results on relative risk and baseline hazard (Gail et al., 1989). The model does not consider many medications such as oral contraceptives, thyroid supplements, rauwolfia preparations, and diazepam medications. Other habits such as cigarette smoking and methylxanthines have not been considered in the analysis. The other major factors that the Gail model does not include are body mass index, age at menopause, breast feeding, plasma estrogen, lobular carcinoma, breast density, second-degree relatives, third-degree relatives, ovarian cancer, and male breast cancer. This model is validated only for white women. The model has been tested in African American women and Asian and Pacific Islander women, and it underestimates the risk in African American women with previous breast biopsies. The data for African American women has been collected from the Women's Contraceptive and Reproductive Experiences (CARE) and from SEER data. The data for Asian and Pacific Islander women has been collected from the Asian American Breast Cancer Study and SEER data. In order to improve the model, various studies have been conducting in Hispanic and other minor community women.

3.2 Claus Model

In the Claus model, the risk is calculated based on an autosomal dominant genetic model. The likelihood of this model is calculated from a joint analysis of mothers and sisters of the

white cases and controls. “The likelihood for the mother of a case was calculated conditional upon the age at onset of the case. For mothers of controls, the likelihood was calculated conditional upon the current age of the control. The likelihood for sisters of cases was calculated conditional upon the breast cancer status of the mother as well as the age at which the case was affected. The age at which a mother with breast cancer was affected, or, in the instance of an unaffected mother, her current age or age at death was also incorporated into the sister’s likelihood. For sisters of controls, the likelihood was calculated in the same fashion with the exception that current age on control was substituted for age at onset of case (Claus, Risch, & Thompson, 1993).

The Claus model is widely used in research studies and clinical counseling. The data used in the Claus model is derived from the Cancer and Steroid Hormone Study. The case-control study is conducted by the Centers of Disease Control. There were 4,730 cases analyzed in the ages of 20 to 54 years, and 4,688 controls were matched to cases on the basis of both geographic regions and five-year categories of age (Claus, Risch, & Thompson, 1991). Family histories are obtained through interviews on different cases about breast cancer. The Claus model considered first- and second-degree relatives and age of onset of relatives. The model incorporates paternal and maternal family history and family history of ovarian cancer. The major limitations of the Claus model are that it may underestimate risk in hereditary families, may not be applicable to all combinations of affected relatives, and does not include risk factors other than family history. The model is best fit for individuals with no more than two first-degree relatives or second-degree relatives with breast cancer.

The Gail model and Clause model underestimate women from hereditary breast cancer families. This model is not a sole model for the families that have the following characteristics (NCI, 2015).

- Three individuals with breast or ovarian cancer diagnosed before the age of 50
- Women who have both breast and ovarian cancer
- People who belong to Ashkenazi Jewish ancestry with at least one case of breast or ovarian cancer

The Claus model assesses the probability of a woman developing breast cancer based on family history. However, it does not consider other risk factors such as bilateral breast cancer, male breast cancer, and ovarian cancer. The Claus model risk can be calculated as lifetime risk or an estimated risk over a ten-year period. The Claus model only considers first- or second-degree relatives, and it does not consider third-degree relatives. In addition, it does not include non-hereditary risk factors.

3.3 BRCAPRO Model

The BRCAPRO model is statistical model associated with software (CancerGene v5.1) to measure the risk of germline deleterious mutation of BRCA1 and BRCA2 genes. This model was developed based on Mendelian approach, and it assumes autosomal dominant inheritance for BRCA1/BRCA2 genes. The model was originally developed in 1995 to 1999 by Parmigiani and colleagues, and it was later modified in 2007. The initial study sample consisted of 292 families that included at least one person tested for BRCA mutation. The family group included 104 African American, 130 Hispanic, 37 Asian American, and 21 other minority families. The families have been identified mainly through two sources. The information for the first group is collected from the Breast Cancer Registry Family (BCFR). The second group was identified by

the Cancer Risk Clinic at the University of Chicago. The information about Ashkenazi families was not included initially and later had been added to the model. The major features that are included in the model are first-degree relatives, second-degree relatives, age of onset of breast cancer, bilateral breast cancer, ovarian cancer, and male breast cancer. This model considers both affected and unaffected relatives. This model gives an estimated result for either BRCA1 or BRCA2 mutation risk. The model computes the breast cancer risk based on the likelihood of BRCA1 or BRCA2 mutation risk. None of the non-hereditary factors such as alcohol intake, age at menarche, oral contraceptive pills use, and breast feeding are incorporated into this model. According to the study done by the Family History Clinic at the University Hospital of South Manchester, this model has the least accuracy in breast cancer risk assessment compared with the Gail, Claus, and Tyrer-Cuzick models (Amir et al., 2003). The model performed best in Hispanics (highest AUC- 0.83) and worst for African American families (AUC-0.68), and it provided intermediate result for other minority groups, mainly Asian-America and Native Americans (AUC-0.71) (Huo et al., 2009). This model appears to be worthy in predicting mutation risk in male patients.

Advantages of BRCAPRO model (Petrucelli, Daly, & Fledman, 2010):

- Estimates probability of mutation in BRCA1/BRCA2
- Considers Ashkenazi Jewish Heritage
- Frequently updated
- Provides printout for pedigree analysis
- Considers both affected and unaffected relatives

Drawbacks of BRCAPRO model:

- Analysis has been done based on high-penetrance families

- Considers only first- and second-degree relatives
- Needs CancerGene software and data entry for each family

The BRCAPRO model performs well in pretest prediction of BRCA mutation, especially in minority communities and Hispanic families, maybe because they are genetically closer than other white populations.

3.4 Cuzick-Tyrer Model

The Cuzick-Tyrer model was designed to predict ten-year risk of breast cancer development using genetic and non-genetic risk factors. The Cuzick-Tyrer model is the most reliable and accurate model for prediction of breast cancer. The Cuzick-Tyrer statistical model calculates the risk of breast cancer by considering many factors such as family history, age at menarche, parity, age at menopause, atypical hyperplasia, age at first childbirth, height, lobular carcinoma in situ, and body mass index. The program assumes that there are other genes that influence breast cancer other than the BRCA genes. The patient's family history is used to estimate the probability of carrying a faulty gene that affects the possibility of developing breast cancer. The data for the model has been collected from the International Breast Intervention Study and other epidemiological data. According to different studies, the Cuzick-Tyrer model is the best model to assess breast cancer risk, as it is integrated with family history, estrogen exposure, and benign breast diseases. The risk factors of the individual are collected through questionnaires and through a review of medical records. A separate analysis has been done for women with atypical hyperplasia, as this disease has a high risk of breast cancer. Among 9,376 patients, 331 women were identified with atypical hyperplasia (Boughey et al., 2010). This model advises genetic counselling when the risk level of BRCA1/BRCA2 is 10% or higher. The IBIS tool (http://www.ems-trials.org/riskevaluator/) is the software tool that has been developed

based on the Cuzick-Tyrer model. The tool does not include risk factors such as lifestyle, breast density, and third-degree relatives. The key benefit over the Claus model and the BRCAPRO model is that the Cuzick-Tyrer model permits the presence of multiple genes of different penetrance levels. The algorithm for this model produces information of BRCA1/2 along with lower penetrance of BRCAX. The model incorporated women who have a single first-degree relative affected by breast cancer. This model also considers ovarian cancer history, as ovarian cancer has a significant effect on breast cancer. The Cuzick-Tyrer model also accurately predicts risk in women whose menarche appeared after the age of 12 years. Even though the model is more powerful than other models at the current time, it is less available to primary care providers than other models.

3.5 BOADICEA Model

The Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) is a statistical model used to compute the risk of hereditary breast and ovarian cancer. This model uses the BRCA1/BRCA2 mutation probabilities to calculate the risk of breast/ovarian cancer. The model was developed based on 2,785 families and has been validated in a large series of families from UK genetics clinics. There are some limitations related to the first version of BOADICEA. "The model assumed that a fixed set of calendar period incidences applied to all cohorts, when breast cancer incidences have been increasing over time. As part of the model-fitting process, we also estimated the *BRCA1* and *BRCA2* breast and ovarian cancer risks, but these were based on a relatively small number of mutation-carrying families and were therefore imprecise" (Antoniou et al., 2008). To calculate the risk, BOADICEA considers the family history of breast, ovarian, prostate, and pancreatic cancer. A new version of BOADICEA has included specific incidents of these cancers from different countries. This model uses the

Mendalian approach to analyze the pedigree of the family history (Lee et al., 2014). The model does not include non-hereditary factors such as age at menopause, alcohol intake, contraceptive pill use, or body mass index. MENDEL is a breast cancer prediction software program that has been developed based on BOADICEA, and it is useful for families of an arbitrary size and structure. The model was developed based on families which are mainly from the UK, and it also included a few families from Sweden. The model was created for invasive breast cancer and is less accurate for ductal cancer in situ (DCIS).

This model groups people who were born in a particular period, and it analyzes the breast and ovarian cancer risk for each person developing cancer during a particular time. The estimated population frequency for the BRCA1 mutation was 0.06% and for the BRCA2 mutation was 0.10% (Antoniou et al., 2008). The model calculates breast cancer risks in BRCA1 and BRCA2 carriers on a birth cohort effect, and the risk increases in more recent birth cohorts. For example, women who born in 1920–1929 show an average cumulative breast cancer risk of 50%, and it is 58% for women born after 1950.

3.6 Comparison of Different Models

The main drawbacks of the Gail model is that it does not consider second- or third-degree relatives to assess the risk and mainly focusing on non-genetic factors. In addition, this model does not count age of onset of breast cancer, bilateral breast cancer, ovarian cancer, and male breast cancer, and it does not calculate the risk of getting breast cancer in the subsequent time. The Gail model measures risk based on patient age, age at menarche, number of prior breast biopsies, age at first live birth, and number of first-degree relatives affected by breast cancer (Evans & Howell, 2007). The tool is designed for health professional such as doctors and nurses.

Therefore, the terms and explanations might not be easy to understand for people who have less knowledge about breast cancer.

Table 2

Comparison of Different Models

Factors considered	Gail model	Claus model	BRCAPRO model	Cuzick–Tyrer model	BOADICEA model
Age (20–70 years)	Yes	Yes	Yes	Yes	Yes
Body mass index	No	No	No	Yes	No
Alcohol intake (0–4 units) daily	No	No	No	No	No
Age at menarche	Yes	No	No	Yes	No
Age at first live birth	Yes	No	No	Yes	No
Age at menopause	No	No	No	Yes	No
Hormone replacement therapy	No	No	No	Yes	No
oral contraceptive pill use	No	No	No	No	No
Breast feeding	No	No	No	No	No
Plasma estrogen	No	No	No	No	No
Breast biopsies	Yes	No	No	Yes	No
Atypical ductal hyperplasia	Yes	No	No	Yes	No
Lobular carcinoma *in situ*	No	No	No	Yes	No
Breast density	No	No	No	No	No
First-degree relatives	Yes	Yes	Yes	Yes	Yes
Second-degree relatives	No	Yes	Yes	Yes	Yes
Third-degree relatives	No	No	No	No	Yes
Age of onset of breast cancer	No	Yes	Yes	Yes	Yes
Bilateral breast cancer	No	No	Yes	Yes	Yes
Ovarian cancer	No	No	Yes	Yes	Yes
Male breast cancer	No	No	Yes	No	Yes

The Cuzick-Tyrer model and the Claus model are developed based on family history that includes first- and second-degree relatives. These tools are not available to the public and are generally made for medical professionals (R. Ellsworth, Decewicz, Shriver, & L. Ellsworth, 2010). Neither of these tools considers male breast cancer history. The Claus model is not linked to important family history factors such as bilateral cancer and ovarian cancer. The Claus and BRCAPRO models were developed fifteen to twenty-five years ago, and worldwide breast cancer incidence has increased by more than 20% since 2008. Therefore, the risk estimates based on lower-risk populations underestimate the current risk. The BOADICEA model is only predicting the risk of BRCA1/BRCA2 mutation (Ellsworth et al., 2010). It does not count other potential genes such as PTEN or Tp53 for the studies.

The Claus and Gail models are mainly used in research studies and clinical risk assessments. The discriminatory accuracy and best fit of the Gail, Claus, BRCAPRO, and Cuzick-Tyrer models have been calculated taking data from 1,933 women attending the Family History and Screening Program in the UK. These models are implemented to estimate the breast cancer risk assessment in women over a follow-up of five years. The ratio of the expected to observed cases of breast cancers was 0.48 for the Gail model, 0.56 for the Claus model, 0.49 for the BRCAPRO model, and 0.81 for the Cuzick-Tyrer model. The accuracy for each model has been developed separately, and the AUC (Area Under the Curve) for the Gail model was 0.735, for the Claus model was 0.716, for the BRCAPRO was 0.737, and for the Cuzick-Tyrer model was 0.762 (Evans & Howell, 2007). The Cuzick-Tyrer model was the most consistent and accurate model. The Claus, Gail, and BRCAPRO models significantly underestimated the risk, especially with people who had a single first-degree relative with breast cancer. The Claus model gives better accuracy if it is manually adjusted for the risk factors. The Cuzick-Tyrer and manual

models were accurate in this case. However, all models accurately predicted the risk in people with multiple relatives diagnosed with breast cancer. The Gail model was less accurate in predicting risk with a single first-degree relative, as it does not count the age at onset of the cancer. The BRCAPRO and Cuzick-Tyrer models were the only accurate models which have approximately similar values to the manual model in perdition of risk with a family history of ovarian cancer. This is the only model which considered ovarian cancer history to predict the breast cancer risk. The Gail, Claus, and BRCAPRO models underestimated the risk in women regarding age of first birth, and these models underestimate the risk factor in women whose menarche appeared after the age of 12 years (Evans & Howell, 2007). The BOADICEA model has almost the same approximate prediction value as the Claus and BRCAPRO models. The BOADICEA model is a polygenetic model, and it examines the risk of BRCA1/BRCA2 mutation along with breast cancer risk. The BRCAPRO and BOADICEA models have high diagnostic accuracy in predicting the mutation risk of BRCA1/BRCA2, and the BOADICEA model has been further modified with tumor markers (ER), progesterone receptors (PR), and human epidermal growth factor receptor 2 (HER2) (Biswas et al., 2012; Tai, Chen, Parmigiani, & Klein, 2008). The limitations of the models are the strength of the research data, the population used, and the overestimation in gene harboring (Aiello-Laws, 2011).

CHAPTER 4

Analyzing Hereditary Information in Predicting Breast Cancer

There are some particular features that may predict the possibility of a hereditary cancer syndrome associated with a single gene mutation. The main features of hereditary cancer syndromes are a family member with more than one cancer, patterns of autosomal dominant transmission, bilateral cancer, rare cancer, precursor lesions, male breast cancer, ovarian cancer, and breast cancer. The information collected can be used to determine whether the cancer in an individual or family is associated with an inherited predisposition and to establish the most likely differential diagnosis (Aiello-Laws, 2011).

The presence of mutated breast cancer genes is detectable in current or future generations based on the information gathered from family history. Family health history plays a major role in determining the risk factors, as it gives a link to hereditary, environmental, cultural, and behavioral factors of the entire family (Yoon et al., 2003). The risk factors involved in breast cancers are age, gender, genetic factors, family history, race/ethnicity, and habits. Breast cancer occurs more often in females than in males because of the female hormones like estrogen and progesterone, which help cancer cell growth (American Cancer Society, 2014). This section analyzes the different genes that are considered for the prediction of breast cancer. The risk of developing breast cancer increases with age. Many factors can be found in a family history that is suggestive of BRCA pathogenic variant. The clues include, multiple individuals on one side of the family with breast or ovarian cancer, a young average age of diagnosis, male breast cancer, associated cancer, individuals with more than one primary cancer, including bilateral breast cancer, Ashkenazi Jewish Ancestry (Murray & Davies, 2013).

4.1 Discussion on Gene Affecting Breast Cancer

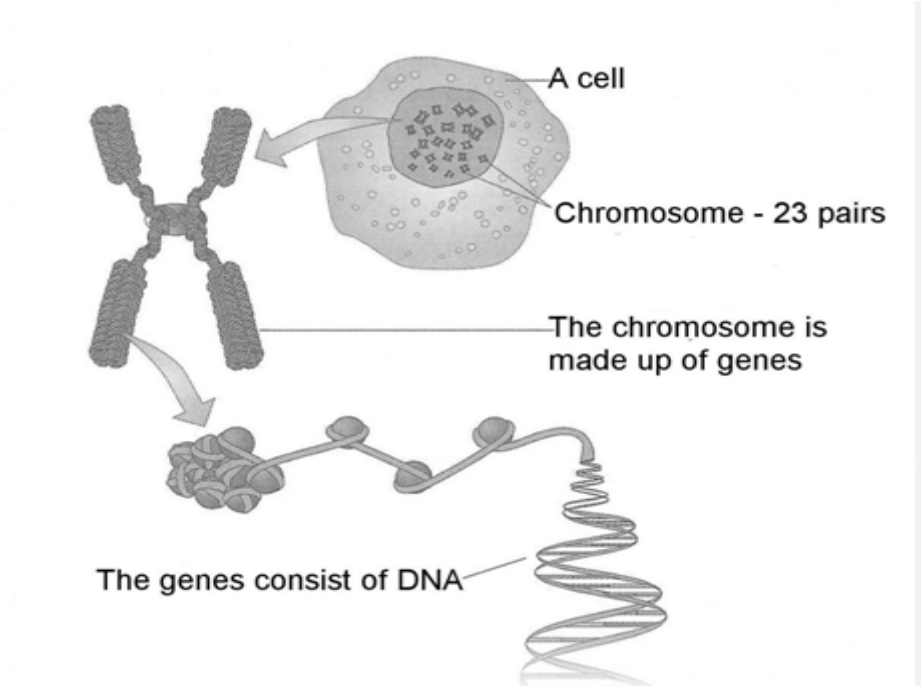

Figure 2. Cell, chromosomes and DNA.

Proteins are large molecule that plays important role in our body. Proteins help the cell to work properly. The proteins are made up with amino acids. There are different kinds of amino acid that together make a protein molecule. The particular sequencing of amino acids make a specific function to each protein. Genes are made up with these different functional molecules (protein) and also it is a part of DNA. A gene is a segment of DNA that is inherited from parents to children, and it determines the traits of the offspring. Every gene has different functions, such as embryo development, cell growth, personality, eye color, or metabolism. DNA is a molecule that stores the genetic information of any living organisms.

DNA is a molecule that contains nucleotides which are the building blocks of DNA. A nucleotide is made up of a phosphate group, a sugar group and a nitrogen base. RNA or

ribonucleic acid is another type of nucleic acid, and it transfers genetic information from DNA to a body cell.

Nucleotides are paired together and form a double helix structure. Each side of the double helix is made up of the phosphate and sugar molecules and the nitrogen bases connect each side. The chains are composed of four different nucleotides, namely Adenine, Guanine, Cytosine and Thymine (Williams, 1997; National Library of Medicine [NLM], 2015). The order of these chemical letters determines DNA's information or genetic information. The nitrogen bases are always found in pairs; adenine pairs with thymine, and guanine pairs with cytosine. Chromosomes are a spiral shaped form of DNA.

The DNA sequence can be mutated in many ways. The different kind of mutation include, missense mutation, nonsense mutation, insertion, deletion, duplication and frameshift mutation. Missense mutation occurs when a DNA base pair sequence change and that substitute one amino acid for another. Nonsense mutation also occurs when the DNA base pair become damaged. In nonsense mutation mutated DNA signals the cell to stop building protein. In insertion mutation, DNA sequencing get altered by another piece of DNA. As a result the gene may not function properly. Deletion mutation changes the DNA sequencing by removing a piece of the DNA sequence. The deleted DNA cause malfunctions in protein synthesis. In duplication a section of DNA is repeated accidently and thus makes the malfunction of the protein. Frameshift mutation occurs when any changes occur in reading frame of DNA. A reading frame include three bases and each base corresponds to one amino acid. When the frameshift mutation occurs it shift these bases and thus amino acid get base that not supposed to have. As a result the protein synthesis may go wrong. A short DNA sequence called Nucleotides are repeating many times in

a DNA sequence row. Trinucleotide repeat 3 times and tetra nucleotide repeat four times. If this chain is repeated unwanted or abnormally makes the protein malfunction.

DNA sequencing can be damaged by many mutagens such as oxidizing agents or alkylating agents and by radiation like X-rays, UV rays, and electromagnetic radiation. Any changes in the DNA sequence of the gene can cause genetic disease. Cancers can be hereditary, and they pass the gene mutations to the next generations through a specific inheritance pattern. Mutated genes such as BRCA1, BRCA2, Tp53, and PTEN confer high lifetime risk of breast cancers (Lalloo & Evans, 2012). There are other genes such as PALB1, BRIP and LKB1 behind breast cancer, which contribute moderate risk. However, BRCA1, BRCA2, TP53, and PTEN are the most considered genes for clinical studies.

Medical information about other factors that cause cancers such as hormone level, gender, lifestyle, diet, habits, pollution, and physical activity helps to predict non-hereditary cancers in the current generation. We estimate the analytical validity of the information that is gathered from the information by comparing it with the results that have been already determined by scientists through experiments.

4.1.1 BRCA1 and BRCA2. Hereditary Breast or Ovarian Cancer Syndrome (HBOC) hereditary genetic condition, and the genes associated with this syndrome are BRCA1 and BRCA2 (BReast CAcer). Most of the breast and ovarian cancers are not hereditary. Both the BRCA1 and BRCA2 genes are responsible for protein involvement in tumor suppression. BRCA1 is located on chromosome 17 and is involved in the repair and regulation of DNA damage. The BRCA2 gene is located on chromosome 13 and is involved in DNA break and repairs (Murray & Davies, 2013). Both genes exhibit high penetrance, ranging from 41% to 90%, with increased risk of other related cancers such as ovarian cancer, epithelial cancer, pancreatic

cancer, and prostate cancer. Male carries of the BRCA gene confer high risk of breast cancer, especially if the gene is BRCA2. The risk of breast cancer increases if the person has close blood relatives who have breast cancer. The risk increases as the number of affected relatives increase and also varies in accordance with people's race and ethnicity. A woman who is an Ashkenazi descendant has more risk of developing breast cancer if she is a carrier of the mutated genes BRCA1 and BRCA2.

It has been calculated that 90% of the early onset breast/ovarian cancer in families is triggered by BRCA1/BRCA2 mutations (Murray & Davies, 2013). Therefore, the risk of developing breast/ovarian cancer is relatively high if the individual has a personal or family history of both cancers. The risk level of developing breast or ovarian cancers due to the mutation of BRCA1/BRCA2 genes may vary even within families with same mutations (Abeliovich et al., 1997; Pagon et al., 2013). The estimation can vary from a 41% to 90% lifetime risk of breast cancer (Antoniou et al., 2003; Chen & Parmigiani, 2007; King, Marks, & Mandell, 2003). A published study shows that the calculated risk of BRCA1 mutation indicates that risk of developing breast and ovarian cancer is 57% and 40%, respectively. In the same scenario with BRCA2, the risk is 49% and 18%, respectively (Chen & Parmigiani, 2007).

Several studies shows that 9% to 21% of BRCA1 mutation is more likely to be linked with triple negative breast cancer (Atchley et al., 2008; Young et al., 2009). In addition to that, BRCA1 mutation appears at early ages among people who have triple negative breast cancer when compared with non-carriers (Gonzalez-Angulo, 2011; Lee, 2011). For a person with triple-negative breast cancer with a family history of breast/ovarian cancer, the risk of BRCA1 mutation was 48% (Fostira et al., 2012). Among Ashkenazi Jewish families, people who have personal histories of triple-negative breast cancer had identified with 11% of BRCA1 mutations.

Different studies show that triple-negative breast cancer is also associated with BRCA2 mutations (Comen et al., 2011). The risk of BRCA2 mutation related to personal history of triple-negative breast cancer ranges from 4% to 17% (Comen et al., 2011). Germline mutations of BRCA1 and BRCA2 mutations associated with prostate cancer, especially BRCA2 mutation, are linked with two to six times' higher risk (Agalliu, Gern, Leanza, & Burk, 2009; Kirchoff, 2004). Patients with prostate cancer due to BRCA2 mutations show decreased survival compared with BRCA1 mutations (Thorne, 2011). BRCA2 mutation is also linked with a higher risk of pancreatic and melanoma cancer than BRCA1 mutation. In addition, BRCA2 mutations have more response with chemotherapy treatment when compared with BRCA1 or non-carrier cases. Germline mutations of BRCA1 and BRCA2 are accountable for 5% to 10% of epithelial ovarian cancer (Jazaeri, 2003). The history of ovarian cancer with BRCA1/BRCA2 mutations is more likely linked to adenocarcinoma. Male carries of BRCA2 mutation also have a high risk of cancer susceptibility. The cumulative lifetime risk for breast cancer among a male who carries BRCA2 mutation is 7% to 8% (Evans et al., 2010). However, men without any mutation have a 0.1% lifetime risk. The NCCN panel recommends that individuals from a family with known BRCA1/BRCA2 mutation should be considered at risk (Evans et al., 2010).

4.1.2 PTEN. PTEN is a tumor suppressor gene, and it regulates normal cell process, including growth, adhesion, migration, invasion, and apoptosis (Yamada & Araki, 2001). Approximately 80% of cancer patients carry a mutated PTEN gene if they meet NCCN clinical criteria for Cowden syndrome. Women identified with Cowden syndrome have high risk of benign fibrocystic breast diseases, and their lifetime risk of breast cancer has been estimated at 25% to 50% with an average age of 38 to 46 years at diagnosis.

Cowden syndrome is considered as very rare disease, and it is underestimated because of the difficulties associated with clinical diagnosis. Cowden syndrome is associated with the tumor suppressor gene PTEN, and it exhibits the autosomal dominant inheritance pattern (Pilarski, 2009). PTEN gene mutation is also related with PTEN Hamartoma tumor syndromes (PHTS), Bannayan Riley Ruvalcaba syndrome (BRRS), Proteus syndrome, Proteus-like syndromes, Lhemitte-Duclos disease, and autism spectrum diseases (Eng, 2014; Pilarski, 2009; Pilarski, Stephens, Noss, Fisher, & Prior, 2011). PTEN mutation is associated with breast cancer, thyroid cancer, endometrial cancer, colorectal cancer, renal cancer, and melanoma. The calculated lifetime risk of cancer is 85% for breast, 34% for renal, 35% for thyroid, 28% for endometrial, 9% for colorectal, and 6% for melanoma (Tan et al., 2012). Women with PHTS have higher risk than men with PHTS. Women diagnosed with Cowden syndrome have high risk of developing fibrocystic breast diseases which has breast cancer risk ranges from 25% to 50%. There are only two cases reported in men with Cowden syndrome (Pilarski, 2009). Approximately 70% people with the PTEN gene mutation are linked with multi-nodular goiter, follicular adenomas, adenomatous nodules, and follicular adenomas. Classic features of Cowden syndrome include mucocutaneous papillomatous papules, trichilemmomas, and palmoplantar keratosis. The study shows that about 40% people with Cowden syndrome have gastrointestinal polyps (hamartomatous, ganglioneuromas). The people who meet the criteria have a risk of carrying the PTEN gene of 85% or higher (Gonzalez et al, 2009; Schaffer, Kamino, Witkiewicz, McNiff, & Orlow, 2006). Early onset colorectal cancer is found in person with PTEN mutations. People with BRRS variant have approximately a 60% risk of PTEN mutation. The people who meet Cowden syndrome criteria will have about an 80% chance of carrying faulty a PTEN gene (Gonzalez et al, 2009). While calculating Cowden syndrome risk, there are some major and

minor features that should be considered. The major features include the personal history of breast cancer, follicular thyroid cancer, endometrial cancer, macrocephaly, multiple gastrointestinal hamartomas or ganglioneuromas, and certain mucocutaneous lesions (Eng, 2000; Pilarski et al., 2009). A person who meets two or more major criteria with one macrocephaly meets the criteria for Cowden syndrome. An individual who meets three or more major criteria is at risk of PTEN mutation. A person with one major criterion and three minor criteria is also considered a person at risk. The minor criteria include colon cancer, autism spectrum disorder, esophageal glycogenic acanthosis, mental retardation, lipomas, thyroid cancer, gastrointestinal hamartoma, and testicular lipomatosis (Eng, 2000; Pilarski et al., 2009). An individual who meets four or more minor criteria can be considered at risk of Cowden syndrome. A first-degree relative with one or more major criteria and two or minor criteria along with a relative diagnosed with Cowden syndrome meets the threshold.

4.1.3 Tp53. Tumor protein Tp53 is another gene that has high penetration to cause breast cancer. Breast cancer is common in women with Li-Fraumeni Syndrome (LFS). There are different tumors associated with LFS syndrome (Gonzalez et al., 2009). The LFS tumor spectrum includes soft-tissue sarcoma, osteosarcoma, pre-menopausal breast cancer, brain tumors, adrenocortical carcinoma (ACC), leukemia, and lung Bronchoalveolar cancer (LFS tumor spectrum). The lifetime risk of breast cancer in a woman with the Tp53 gene is 49% by the age of 60. According to the clinical study, a patient with invasive breast cancer and a family history of no core cancer has a 0% chance of having the p53 mutation. However, a person with breast cancer under the age of 30 and with a family history of one or more core cancers in first- or second-degree relative has a 100% chance of having the p53 mutation. Patients younger than 30

with unilateral breast cancer and no cancer in first- or second-degree relative have a 7% chance having the p53 mutation (Gonzalez et al., 2009).

Li-Fraumeni syndrome is a rare hereditary diseases associated with germline mutations of Tp53. Only 1% of hereditary breast cancers are related with germline mutation of Tp53. It is located on chromosome 17 and it is known as "guardian of genome." Tp53 gene plays an important role in cell cycle and apoptosis. An individual who meet Classic Li-Fraumeni syndrome criteria has a 50% - 70% risk of Tp53 gene mutation. LFS is a high penetrant cancer syndrome, and it is associated with soft-tissue sarcomas, osteosarcomas, acute leukemia, colon cancer, premenopausal breast cancer, adrenal cortex cancer, and brain tumors. Sarcoma, breast cancer, adrenocortical tumors, and brain tumors are known as core cancers of LFS syndrome, as they are mostly related to Tp53 mutation. Studies show that HER-positive breast tumors are highly associated with Tp53 mutation. Individuals with soft-tissue sarcomas, brain tumors, or adrenocortical carcinomas at early ages will have a high lifetime risk of developing multiple primary cancers. When compared with males, female breast cancer is more often related to LFS syndrome. NCCN guidelines give different criteria to identify LFS syndrome. Classic LFS syndrome criteria include individuals with known Tp53 mutation, individuals diagnosed with sarcoma before the age of 45 years, a first-degree relative diagnosed with cancer before the age of 45, and an additional first- or second-degree relative with the same cancer diagnosed at the age of 45 or younger or with sarcoma at any age. Classic LFS has high prediction value (56%) and specificity. Some other works have been done by different groups to make the LFS criteria more accurate and predictive. Birch and colleagues have added more cancers to the LFS criteria. Chompret and his colleagues added many criteria, including patients with multiple tumors, at least two core tumor types diagnosed before the age of 36, and patients with adrenocortical

carcinoma diagnosed at any age. The Chompret criteria has 20% to 35% prediction values, and when it is added to classic LFS criteria, it improves the prediction value to 95%. Women who identify breast cancer before the age 35 with or without family history of core cancers of LFS syndrome can be considered at risk of Tp53 mutation. Studies show that women with breast cancer before the age of 35 have a 3% to 8% risk of Tp53 mutation. A member from a family with known deleterious Tp53 mutation is always at risk even without any symptoms of cancer.

4.2 Inheritance Patterns of BRCA1, BRCA2, Tp53 and PTEN Genes

Humans have twenty-three pairs of chromosomes. The first twenty-two are called AUTOSOMES and the twenty-third pair is the SEX CHROMOSOMES (the pair of chromosomes that determines the gender of the offspring) (NLM, 2015). Mutated genes are inherited from biological parents in specific ways. Autosomal dominant inheritance is one of the basic patterns of inheritance. Autosomal dominant inheritance means that the gene carrying a mutation is located on one of the autosomes (chromosome pairs 1 through 22). This means that males and females are equally likely to inherit the mutation. In autosomal dominant mutation, just one of the two copies of a particular gene is enough for a person to have a trait, such as an increased risk of developing cancer (Chapman, 2007). Autosomal recessive inheritance means that the gene carrying the mutation is located on one of the autosomes (chromosome pairs 1 through 22). In autosomal recessive inheritance, both copies of the gene must have a mutation in order to make a person have the trait. A person who has only one recessive gene mutation is said to be a "carrier" of the trait or disease, but he/she does not have any health problems from carrying this one mutated gene (Williams, 1997).

Genetic conditions are caused by mutations in single genes and can be inherited in different ways. Single gene inheritance is called Mendelian inheritance, and there are four

different types of Mendelian inheritance patterns: 1) autosomal dominant, 2) autosomal recessive, 3) X-linked recessive, and 4) X-linked dominant (National Library of Medicine, 2015).

4.2.1 Autosomal dominant inheritance. Autosomal dominant inheritance means that the gene carrying a mutation is located on one of the autosomes (chromosome pairs 1 through 22). In autosomal dominant inheritance, only one copy of the faulty gene is sufficient to make the individual at the risk of the disease. Autosomal dominant disorders are likely to occur in each generation of an affected family.

When a parent has a dominant gene mutation, there is a 50% chance that any child he/she has will also inherit the mutation. There are four possible combinations in the children. Two of the four, or 50%, have inherited the mutation. The other 50% have not inherited the mutation. These four combinations are possible every time a pregnancy occurs between these two individuals. The gender of the children (whether they are sons or daughters) does not matter. The chance is fifty-fifty for each pregnancy. BRCA1, BRCA2, PTEN and Tp53 genes exhibit the autosomal dominant inheritance pattern.

4.2.2 Autosomal recessive inheritance. In each cell, two mutated copies of the gene should be present to make the person at risk of disease. "Recessive" means that both copies of the faulty gene should be inherited to make the person susceptible. A person who has only one recessive gene mutation is said to be a "carrier" for the trait or disease, but he/she does not have any health problems from carrying this one mutation. Most people do not know they carry a recessive gene mutation for a disease until they have a child with the disease. Once parents have had a child with a recessive disease, there is a one out of four, or 25%, chance with each subsequent pregnancy for another child to be born with the same disorder. This means that there

is a three out of four, or 75%, chance for another child to not have the disease (National Library of Medicine, 2015).

4.2.3 X-linked dominant inheritance. X-linked dominant conditions are affected by mutations in genes on the X chromosome. Females are more commonly affected than males, and passing on an X-linked dominant disorder has a different outcome for men and women. Families with X-linked dominant diseases have frequently affected both males and females in each generation. In X-linked inheritance, the father cannot pass the faulty gene to their sons (NLM, 2015).

4.2.4 X-linked recessive inheritance. These disorders are due to the mutations in genes on the X chromosome. Males are more often affected than females. X-linked recessive disorders are frequently found in males but rarely affect females.

The next section will explain how the inheritance pattern is used for this software tool. The BRCA1/BRCA2, PTEN, and Tp53 genes exhibit autosomal dominant inheritance pattern. The mutated gene can be inherited by the next generation in different ways, and it depends on the gene that is involved with the disease. Consider a mother who is the carrier of the abnormal gene "a" and the father who is carrying the normal gene "A", as shown in Figure 3. In hereditary breast cancer, the gene passes to the next generation through an autosomal dominant inheritance pattern. Every child inherits two copies from their parents, one from their mother and one from the father (National Center for Biotechnology Information [NCBI], 2009; HHS, 2013). When a parent has a dominant gene mutation, there is a 50% chance that any child from each pregnancy will inherit the mutation. The BRCA1, BRCA2, Tp53, and PTEN genes exhibit the autosomal inheritance pattern.

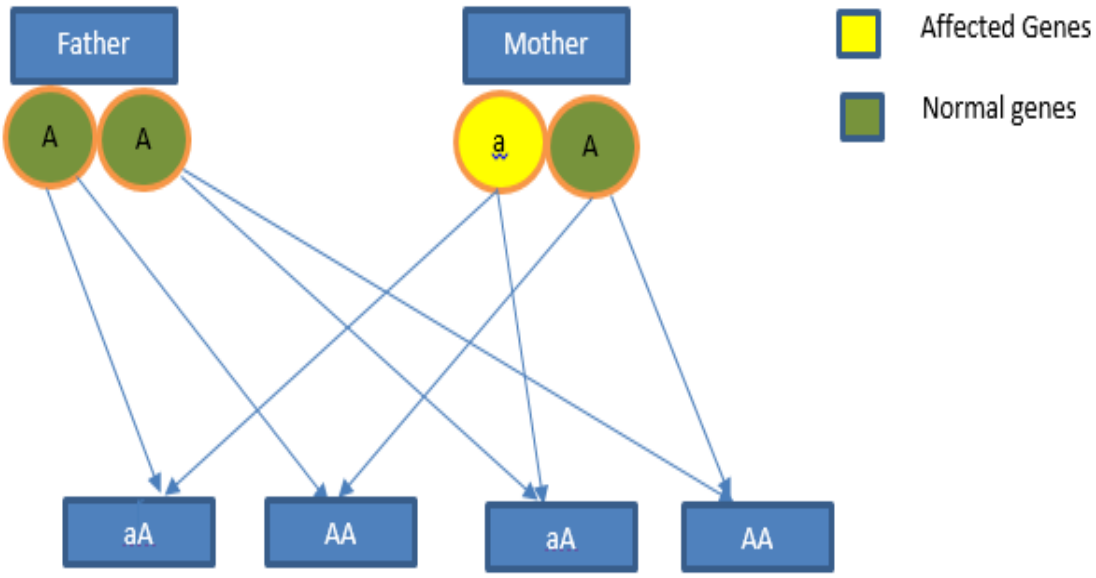

Figure 3. Autosomal dominant inheritance pattern.

A person with autosomal dominant inherited diseases will have four different possibilities with each pregnancy. Two out of the four, or 50%, can inherit the mutated genes (aA, aA). The other 50% do not inherit the mutated genes (AA, AA). These four combinations are possible every time a pregnancy occurs. The gender of the children does not matter.

4.3 How the Traits are Passing Down to the Generations?

Most cancers are non-hereditary, and hereditary cancers are commonly caused by genetic defect. There are two main categories of mutation based on the type of cell that the genetic change occurs in. Mutations that are passed from parents through the egg or sperm are known are inherited mutations. Somatic mutations can happen in the lifetime of any person. They are not inherited from their parents and it will not pass to the next generations.

Two main types of genes that play a role in cancer are proto-oncogenes and tumor suppressor genes. Tumor suppressor genes control cell division and repair DNA mistakes. When tumor suppressor genes don't work properly, cells can grow out of control, which can lead to

cancer. Many different tumor suppressor genes have been found, including TP53 (p53), BRCA1, BRCA2, APC, and RB1. Proto-oncogenes speed up the cell division and in any mutation can cause uncontrollable growth of the cells. The mutated proto-oncogenes are called oncogenes. When this happens, the cells grow out of control, which can lead to cancer.

The mutated gene can be inherited to the next generation in different ways, and it depends on the gene that involved with the disease. Genes are found in pairs. Each parent contributes one copy of the gene to their child. These gene copies can be inherited from parents in different ways, which are called inheritance patterns.

For example, some of the hereditary disease shows autosomal recessive inheritance. That means two copies of the abnormal genes must be present in order for the disease or trait to develop. Most people do not know that they carry a recessive gene mutation for a disease until they have a child with the disease. Once the mother has conceived a child, there is 25% chance of getting a normal child without the mutated gene. There will be a 25% of chance a child born with the disease. This means that there is a two out of four, or 50%, chance of getting a child who carries the mutated gene and may not show any symptoms of the disease in his or her lifetime.

CHAPTER 5

Prediction of Hereditary Breast Cancer

The capabilities of today's digital world make innovative progress in health care. Doctors take advantage of fully functional and exchangeable electronic health records to provide better treatment and care. Health records that are stored in an electronic format deliver high accuracy in diagnoses and health outcomes. Nowadays, people use the Internet and medical websites to get more health information about diseases. Cancer is predictable in current or future generations based on the information gathered from family history tools and genetic and medical information. Collecting data from close relatives may be sufficient in determining the risk. However, risk prediction will be limited if the families are small. Predictions of diseases like breast, ovarian, colorectal, and prostate cancers have high degrees of accuracy if they have been developed based on family history (Santaguida et al., 2007). Identifying repeated diseases and associated genes would help to forecast the disease in the future generation (offspring). As we are concentrating on breast cancer, the next section will explain how a family history will be useful in predicting breast cancer in any individual or their next offspring. The aim of this research is to develop a software tool to predict hereditary breast cancer. We combine the medical, genetic, and family health history reports to create the program.

5.1 Flowchart to Predict the Hereditary Breast Cancer Risk

There are different models and software programs that have been developed to predict the risk of breast cancer (Evans & Howell, 2007). However, these tools do not consider other genes such as TP53 and PTEN that contribute a high risk for breast cancer. In this computer program, we have combined BRCA1/BRCA2, TP53, and PTEN gene risks. We are using National Comprehensive Cancer Network guidelines to predict BRCA1, BRCA2, TP53, and PTEN

mutation risk. According NCCN members Daly et al, the guidelines for breast and ovarian cancer are focusing on assessment of mutations in the genes BRCA1, BRCA2, TP53, and PTEN, and recommended approaches to genetic testing and management strategies.

The National Comprehensive Cancer Network (NCCN) is an association of twenty-five global leading cancer centers. The association develops guidelines for most cancers and is updated by forty-seven individual panels, including over 950 clinicians and oncology researchers from the twenty-five institutions that are members of the NCCN. The panel members include clinicians and researchers, and they have several academic disciplines. According to NCCN guidelines (Daly et al., 2014), "It should be emphasized that these guidelines were not developed as a substitute for professional genetic counseling. Rather they are intended to serve as a resource for healthcare providers to identify individuals who may benefit from cancer risk assessment and genetic counselling, to provide genetic counselors with an updated tool for the assessment of individual breast cancer and ovarian cancer risk and to guide decisions related to genetic testing, and to facilitate a multidisciplinary approach in the management of individuals at increased risk of hereditary breast and ovarian cancer. Although cancers other than breast and ovarian cancers are associated with these hereditary syndromes, the main focus of this NCCN Guidelines is on the management of breast and ovarian cancer risk in these individuals" (Daly et al., 2014). Moreover, new guidelines go through annual reviews, and it will be circulated among multidisciplinary researchers for comments at each NCCN branches.

The flowcharts have been developed based on NCCN guidelines, and the analysis has been divided into three sessions: 1) Hereditary breast/ovarian cancer syndrome analysis 2) Li-Fraumeni syndrome analysis and 3) Cowden syndrome analysis. These three syndromes are related to germline mutation of BRCA1/BRCA2, TP53, and PTEN genes that have high impact

on hereditary breast cancer diagnoses. The below figures give details about the logical flow of the program. In this program, we have used the latest version of the guidelines (Daly et al., 2014).

5.1.1 Checking BRCA1/BRCA2 mutation.

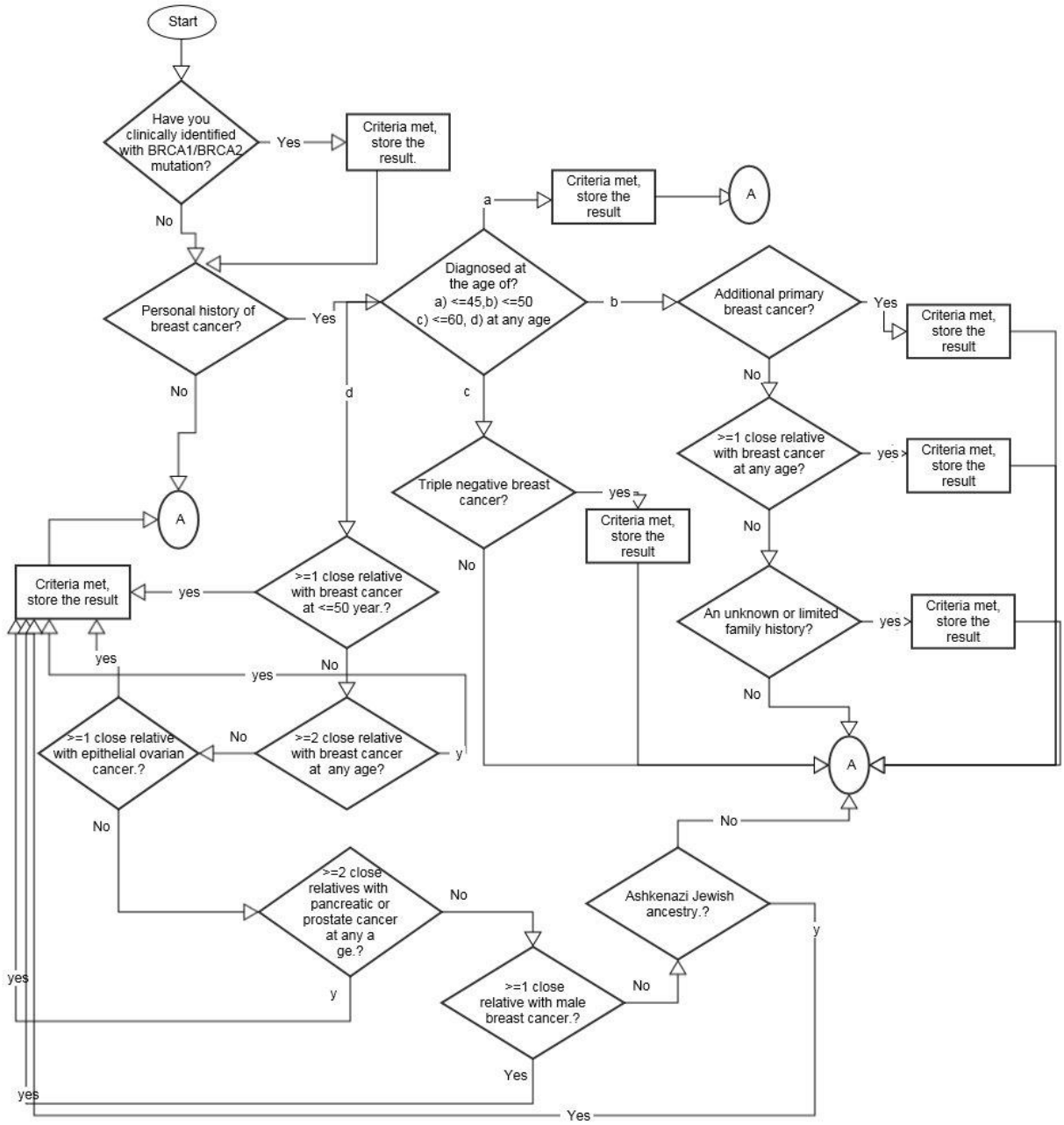

Figure 4. BRCA1/BRCA2 personal history analysis.

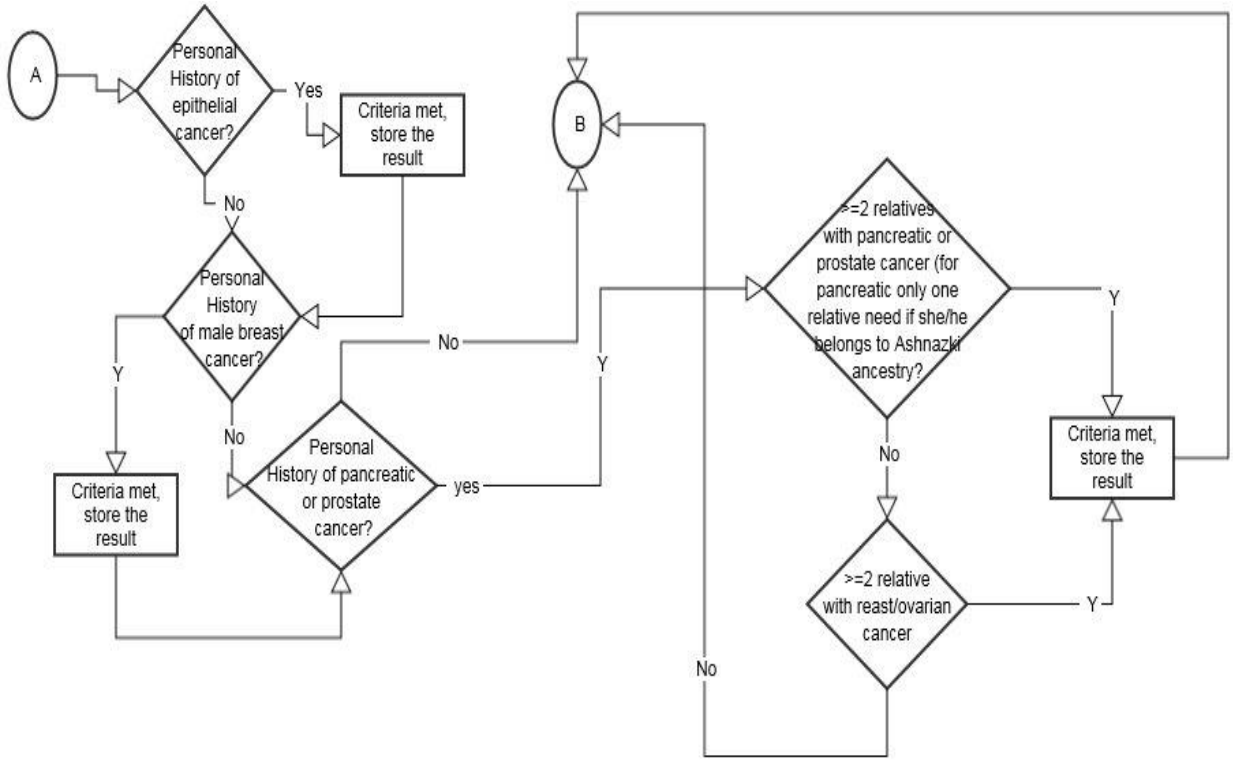

Figure 5. BRCA1/BRCA2 personal history analysis (cont.).

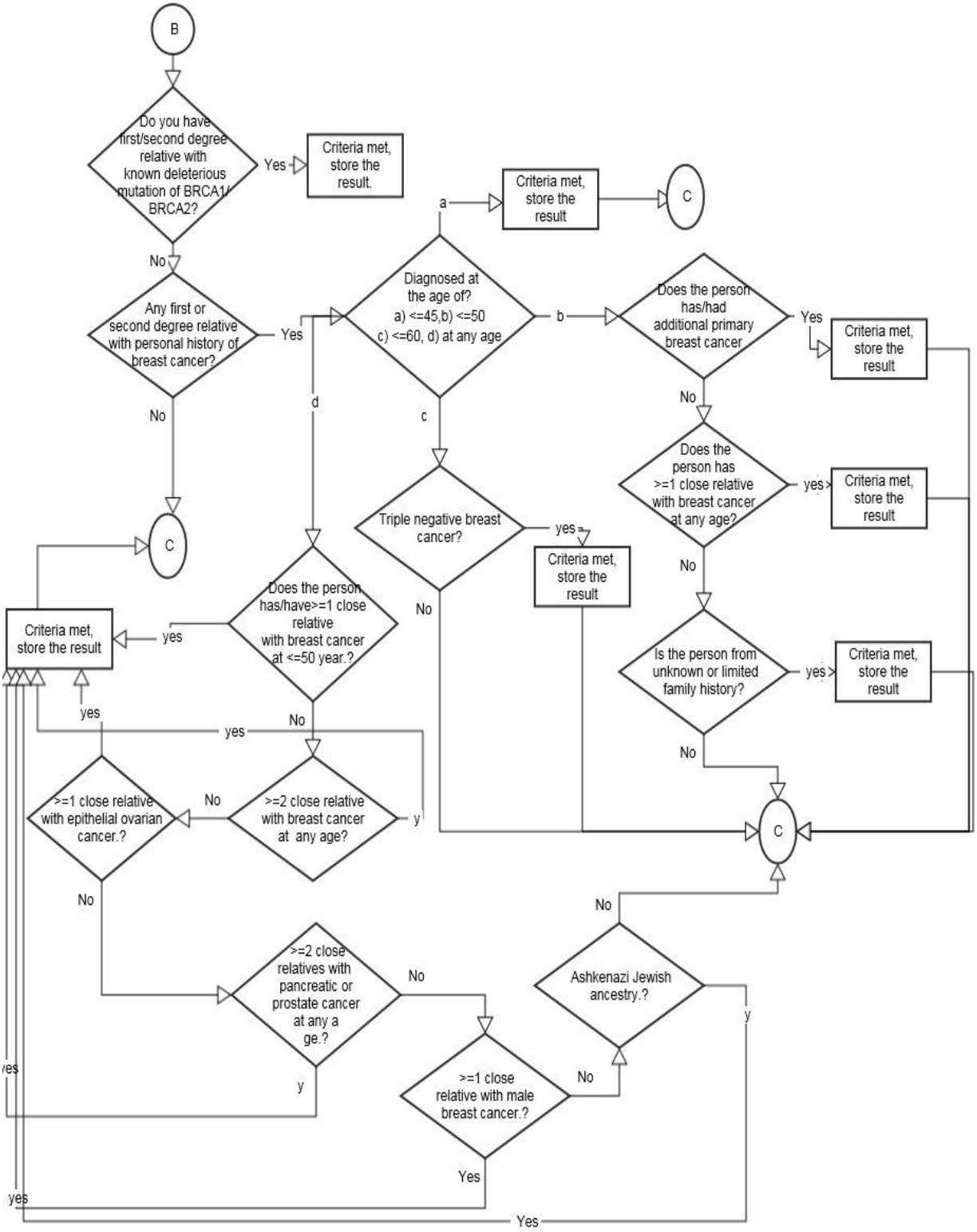

Figure 6. BRCA1/BRCA2 family history analysis.

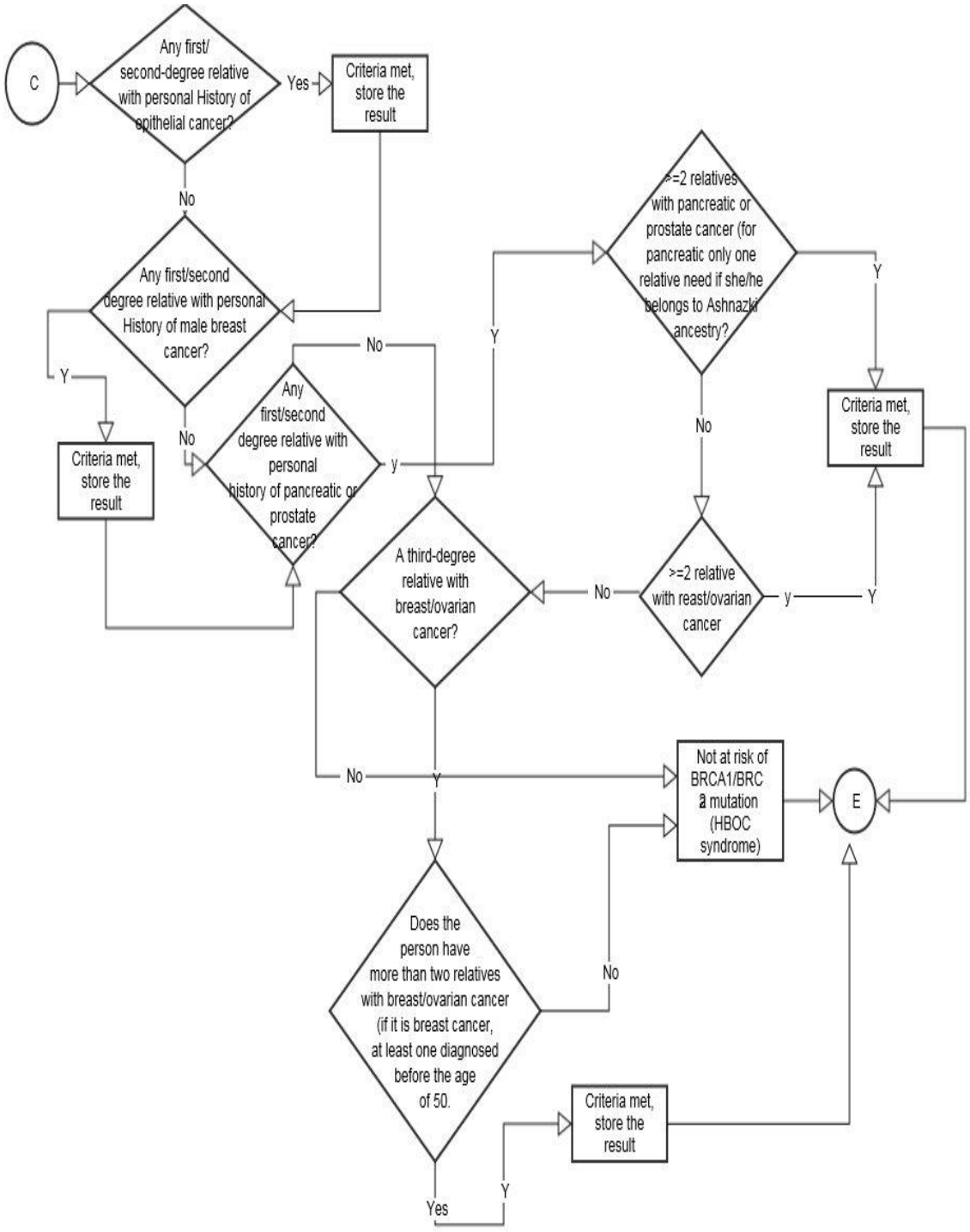

Figure 7. BRCA1/BRCA2 family history analysis (cont.).

The figure 4 and figure 5 shows the procedures of personal history analysis of BRCA1 or BRCA2 mutation risk. After collecting the information about already known or diagnosed mutation of BRCA1/BRCA2 gene, it checks the personal history of breast cancer.

In order to check the personal history of breast cancer, the flowchart has been divided into four groups based on different age category. The risk criteria will be met if the person is under the age category ‘a’ (age <=45), and there are no additional criteria needed to satisfy the ‘at risk’ condition. The criteria will be met, if the person is under or at the age of 50, and if she/he has any additional primary breast cancer or more than one close blood relatives with breast cancer. The person will be at risk if she/or he has an unknown family history. If the age category is 60 or under, the result will be stored if the person is identified with triple negative breast cancer. For ‘at any age’ category, and the criteria will be met if more than one close blood relatives with breast cancer at 50 or younger age, or more than two close blood relatives with breast cancer at any age, or more than one close blood relatives with male breast cancer, or more than close blood relatives with pancreatic or prostate cancer. The result will be stored, if the person belongs to Ashkenazi Jewish Ancestry, and if the person has a personal history or family history of breast or prostate cancer.

After that, the result will be stored if the person has an epithelial ovarian cancer history. In addition, the flowchart checks the personal history of male breast cancer. Next, it checks the personal history of pancreatic or prostate cancer.

After checking the personal history of pancreatic or prostate cancer history, the family history of breast, ovarian, pancreatic and prostate cancers will be checked. If there are more than two relatives from the same side of the family is affected with any of these cancers, the result

will be stored as ‘at risk’. For pancreatic cancer, if the person belongs to Ashkenazi Ancestry, then only one additional affected relative is needed to meet the criteria.

The figures 6 and 7 show the procedures of family history analysis. The same steps that has been used for personal history analysis will be used to check the family history (only for first- or second-degree relatives) analysis.

After collecting the information of already known or diagnosed gene mutation, the relative’s family history will be checked. For that, the flowchart collects the information about first- or second-degree relative’s personal history of breast cancer. Checking personal history of breast cancer has divided into four based on the age group. The information will be collected separately for different age category.

After that, the flowchart checks if the person has a history of epithelial ovarian cancer. Next, it will check the presence of pancreatic and prostate cancer and also checks the male breast cancer history. If the person has a first- or second-degree relative with these cancers, the result will be stored and enter into next section.

To check a third-degree relative’s family history, a different method has been followed. If a third-degree relative is identified with breast or ovarian cancer, the flowchart checks the number of other relatives who have identified with breast or ovarian cancer in the same side of the family. If there are more than two people identified with breast or ovarian cancer (one at least before the age of 50), the result will be stored as ‘at risk’.

5.1.2 Checking Tp53 mutation.

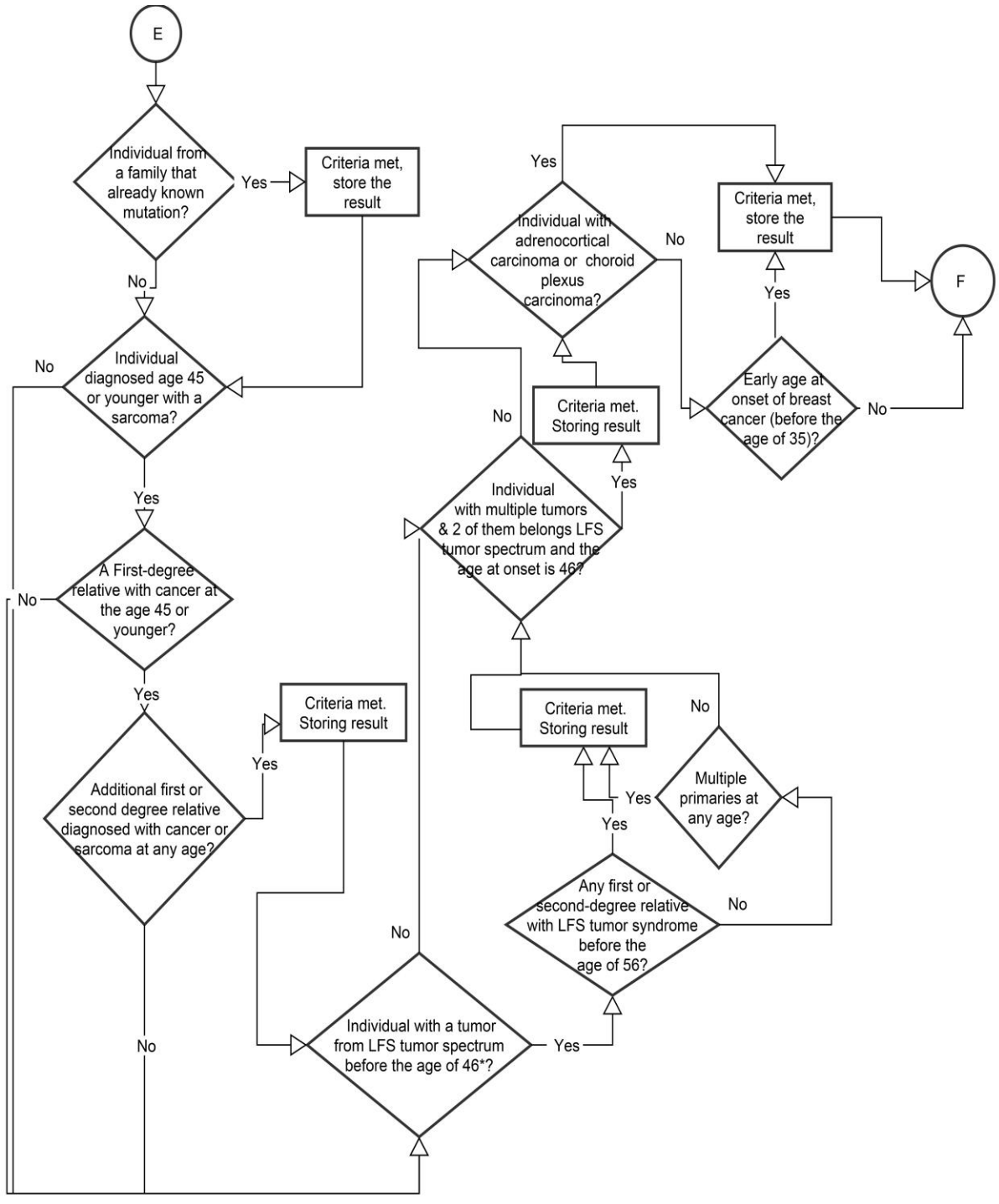

Figure 8. Tp53 mutation analysis.

The figure 8 shows the procedures to check the Tp53 mutation. To begin with, the information will be collected if the person is already identified with Tp53 mutation.

In the next step, it checks whether the person meets classic-Li-Fraumeni syndrome criteria or not. In order to check the classic Li-Fraumeni syndrome criteria, the flowchart checks if the person is identified with a sarcoma before the age of 45. If the person is diagnosed with sarcoma, first- or second relative's history of cancer before the age of 50 will be checked. If the first-, or second-degree relative is identified with cancer, it checks the additional relative's history of cancer. The result will be stored as 'at risk' if all three criteria has been met.

Next, the procedure enters into Chompret criteria section. Checking Chompret criteria starts from identifying individual with a tumor that belongs to LFS tumor spectrum before the age of 46. LFS tumor spectrum includes sarcoma, brain tumor, breast cancer, leukemia, adrenocortical carcinoma and lung Broncho alveolar cancer. Next, the relative's history of LFS tumors (include breast primaries) will be checked if the individual is identified with LFS tumors. After that, it checks whether the individual is diagnosed with multiple tumors, and two of them belongs to LFS spectrum or not. If yes, the result will be stored as 'at risk'. In the final step, the flowchart checks if the individual is diagnosed with adrenocortical carcinoma or Choroid plexus carcinoma.

In the next step, the flowchart checks if the person is diagnosed with breast cancer before or at the age of 35-year-old. If the person meets the criteria, he/she is at risk of both Tp53 and BRCA gene mutation. The result will be collected if the condition is satisfied, and enter into PTEN mutation analysis.

5.1.3 Checking PTEN mutation.

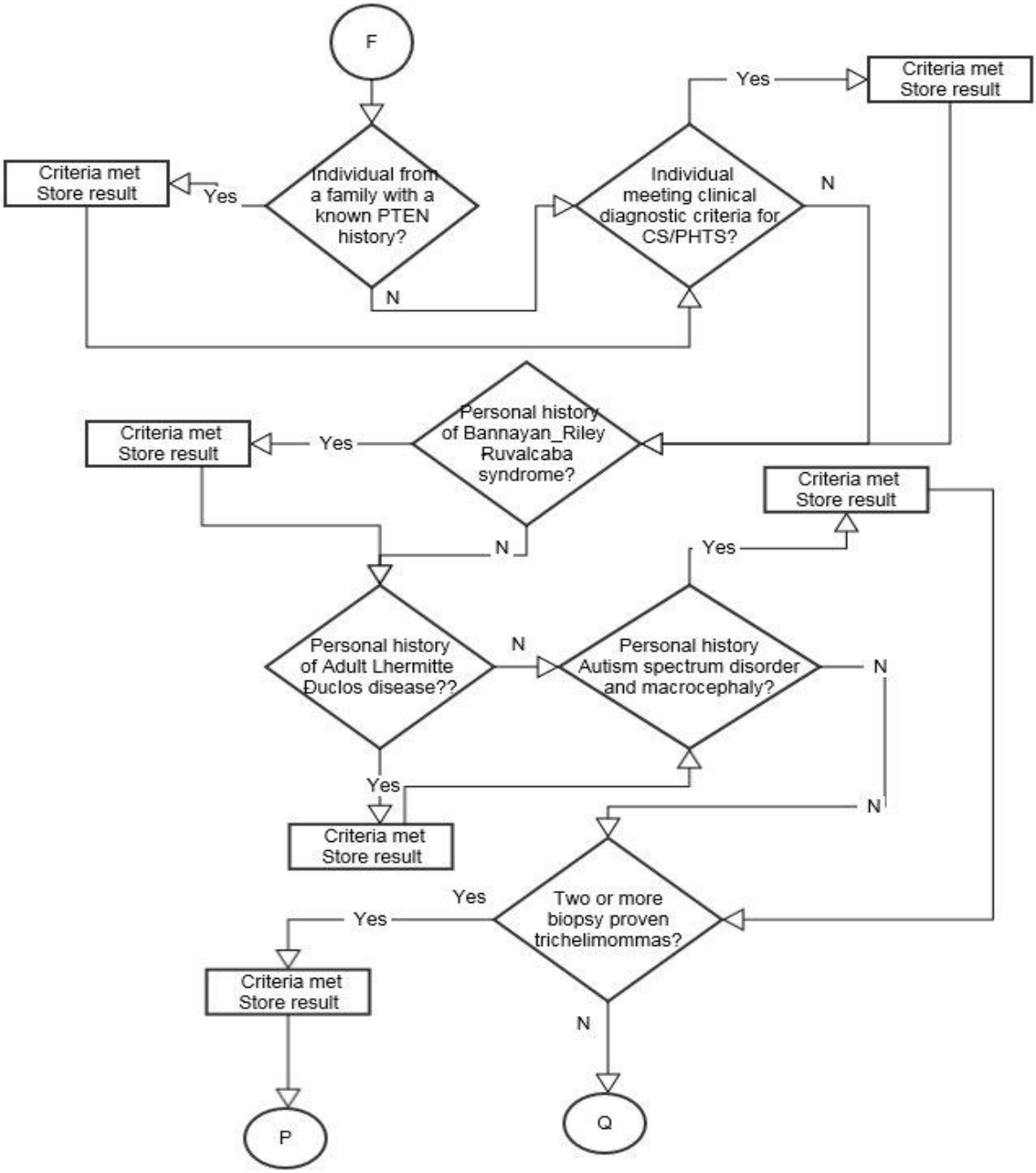

Figure 9. PTEN analysis.

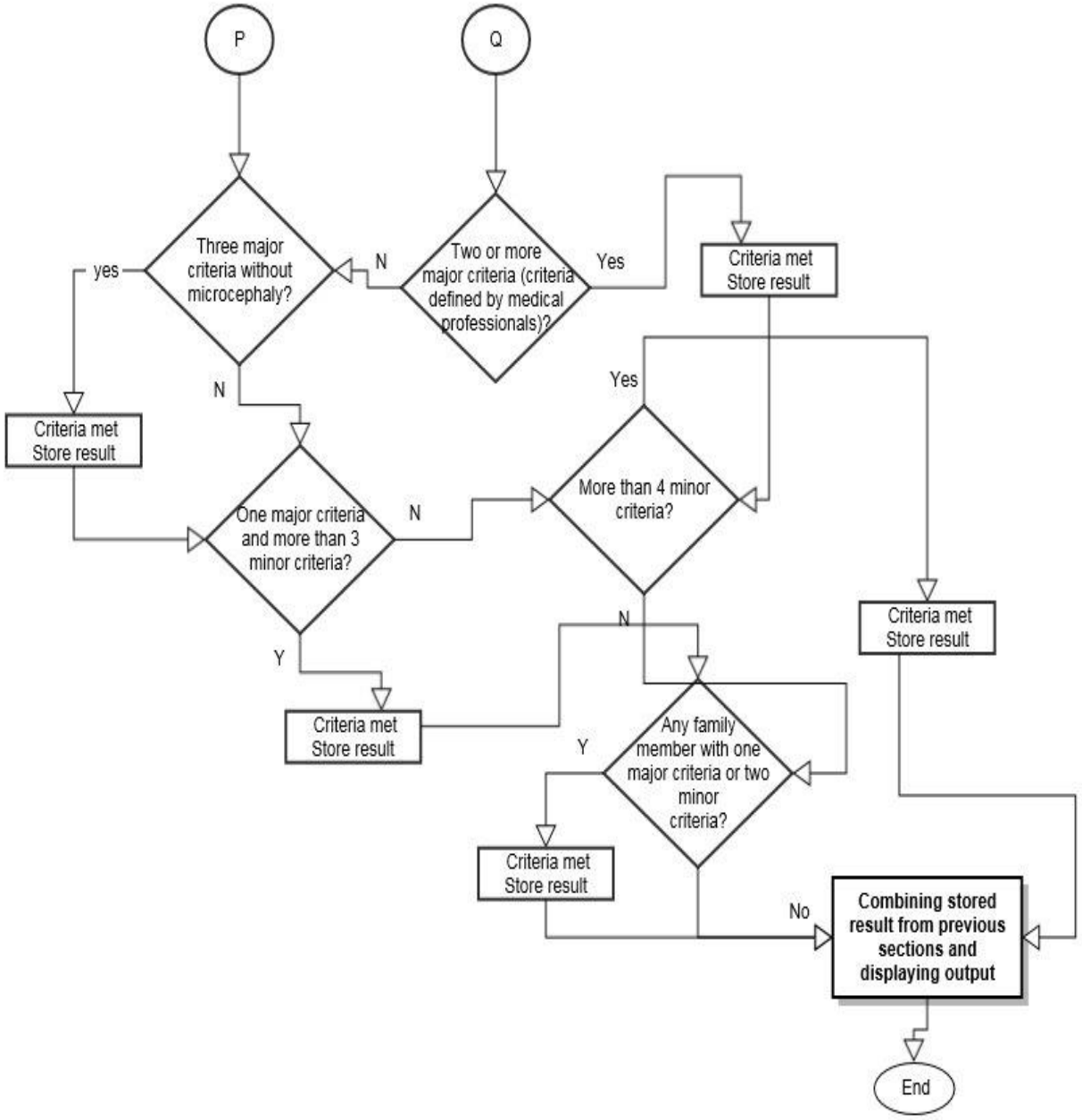

Figure 10. PTEN analysis (cont.).

The figures 9 and 10 explain the procedures involved in PTEN mutation analysis. First of all, the flowchart checks if the person is already known or diagnosed PTEN gene mutation, and then the result will be collected. In next step, it checks if the individual is meeting clinical

diagnostic criteria for CS/PHTS and the information will be collected if the person has met the criteria. After that, it checks the personal history of Bannayan_Riley Ruvalcaba syndrome, and the information will be stored if it is identified. Next, the flowchart enter into checking of personal history of Adult Lhermitte Duclos disease, and the result will be stored if the person has/had a history of the disease. Then, it checks the personal history of autism spectrum disorder and macrocephaly condition. After that, it checks two or more biopsy-proven trichilemmomas. After collecting the result, it checks whether the person is meeting three major clinical diagnostic criteria without macrocephaly. Next, the flowchart enters into identification of two or more major criteria with macrocephaly or it checks whether the person has met one major criterion and more than three minor criteria. In the next step, the flowchart checks whether the individual meets more than four minor criteria or not. After that, the flowchart checks if any of the individual's family member meets one major criterion and two or more minor criteria. The result will be stored if the criteria have been met. In final section, the result stored from different sections will be combined and displays the result based on the different combination of the result. The final output gives a separate report based on the three sections.

5.2 Proposed Design

The proposed block diagram of the program is shown in Figure 4. We have considered all the clinical possibilities, including inheritance pattern, family history assessments, and medical reports to generate the program.

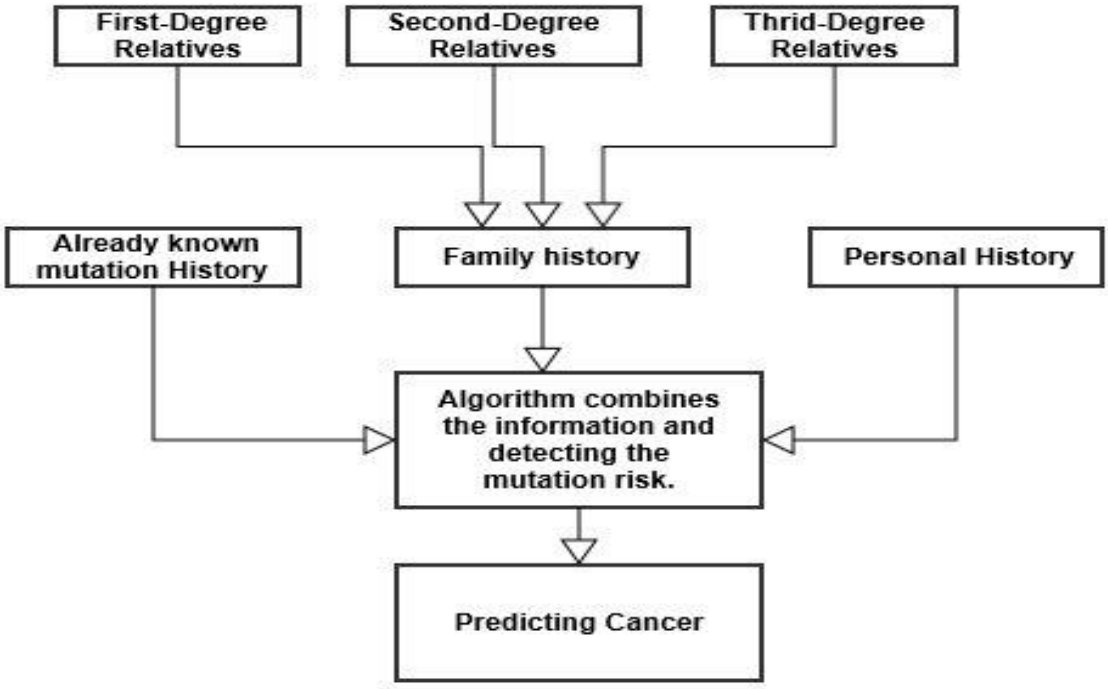

Figure 11. Proposed design.

As described earlier, this program has three sections (BRCA1/BRCA2 checking, Tp53 checking and PTEN mutation checking). As shown in the above figure, the program analyzes personal history, family history and already known mutation condition. The program checks individual known mutation history, personal history, and family history for BRCA1/BRCA2, Tp53 and PTEN gene mutation. The result of each section will be stored separately, whenever the NCCN criteria are met. The program combines different section's result and displays the result based on the analysis. The result will contain the risk of passing mutation risk or getting cancer in next generation. The family history includes, first, second and third-degree relatives health history.

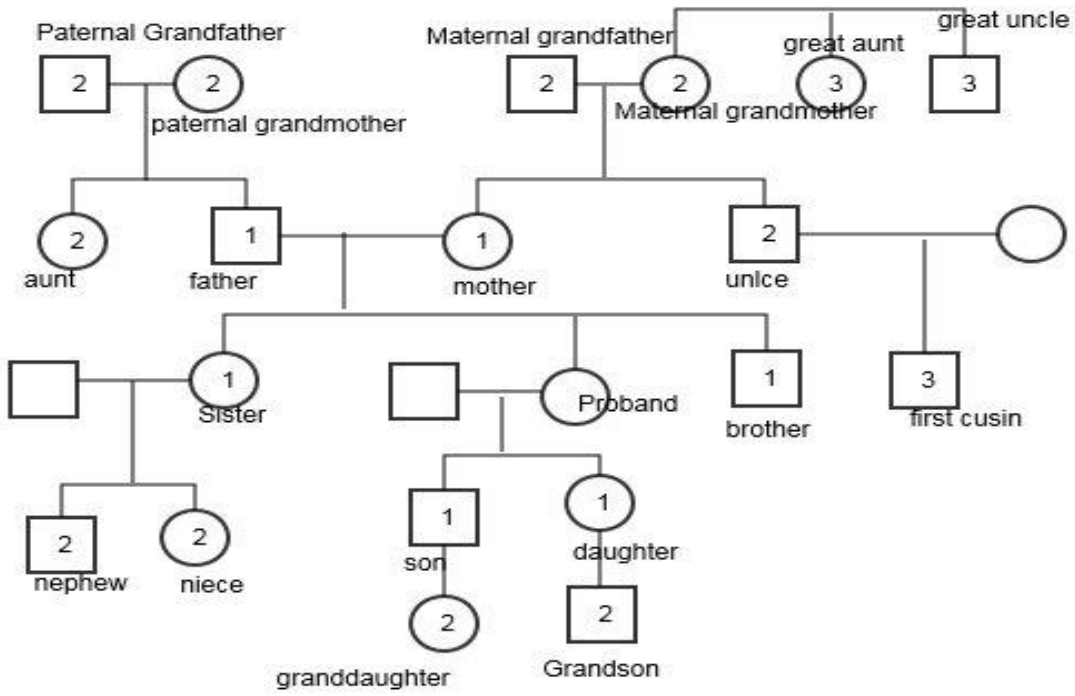

Figure 12. A pedigree chart.

We have considered family history that includes first-, second-, and third-degree relatives. The above figure 6 gives a detailed picture of first-, second-, and third-degree relatives. First-degree relatives include parents, siblings, and children. Second-degree relatives include grandparents, aunts, uncles, nieces, nephews, grandchildren, and half-siblings. Third-degree relatives include great-grandparents, great-aunts, great-uncles, great-grandchildren, and first cousins (Daly et al., 2014). The tool combines the personal history, family history, and known mutation history to detect the presence of faulty genes. In the output, the tool combines all the results from different sections and displays the results.

5.3 Software Tool to Predict Hereditary Breast Cancer in Current and Future Generation

This software program was developed with JavaScript, HTML, and CSS.

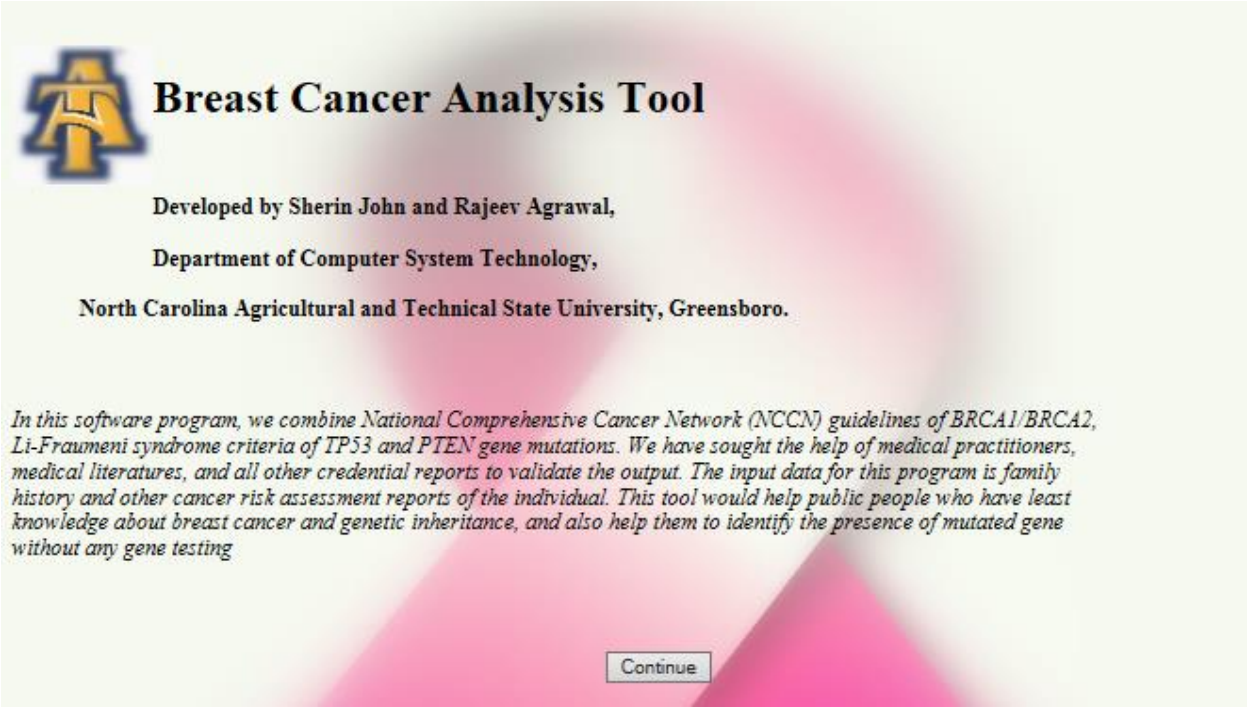

Figure 13. Image of the software program.

Based on the guidelines, the program analysis is divided into three sessions: 1) Hereditary breast/ovarian cancer syndrome analysis, 2) Li-Fraumeni syndrome analysis, and 3) Cowden syndrome analysis.

5.3.1 Hereditary breast and ovarian cancer syndrome analysis. As the first step, the program checks the breast and ovarian cancer syndrome (BRCA1/BRCA2 mutation). This part of the program would check different criteria that are defined by NCCN to detect the presence of BRCA1/BRCA2 genes. Based on the NCCN guidelines, individuals from a family with known BRCA1 or BRCA2 mutation should be considered for testing. If the individual is not from a known gene mutation, the risk can check based on the testing criteria discussed below. Close relatives include first-, second- and third-degree relatives from the same side (paternal and

maternal side) of the family. If the individual has less than two first- or second-degree relatives can be considered as limited family history. The analysis has been divided into three sections: 1) an individual with an already known or diagnosed BRCA1/BRCA2 mutation, 2) checking personal history of different types of cancers, such as breast cancer, epithelial ovarian cancer, male breast cancer, pancreatic or prostate cancer, and Ashkenazi ancestry, and 3) checking family history of the person. The program checks personal and family histories to detect the presence of mutation. The tool also checks the presence of different hereditary cancers such as epithelial ovarian cancer, prostate cancer, pancreatic cancer, and male breast cancer, and it also checks ancestry (e.g., Ashkenazi Ancestry). The result will be stored if any criteria are met.

5.3.1.1 Person with known mutation. This part will check whether the individual is already diagnosed with BRCA mutations or if he or she has a known personal/family history of BRCA1/BRCA2 mutation. If the person has a genetic report or medical report that identify the mutation of BRCA1/BRCA2 gene, she/he will be considered as at risk of developing breast cancer. The lifetime risk vary from 41% to 85%. If there is a history of already known BRCA1/BRCA2 mutation, the tool will consider the person as having met the NCCN criteria and will store the result as positive.

5.3.1.2 Personal history. In the next section, the program analyzes the personal history of cancers that related with BRCA1 and BRCA2 mutation. Checking personal history has been divided into four subsections: a) personal history of breast cancer, b) epithelial ovarian cancer, c) male breast cancer and d) pancreatic or prostate cancer.

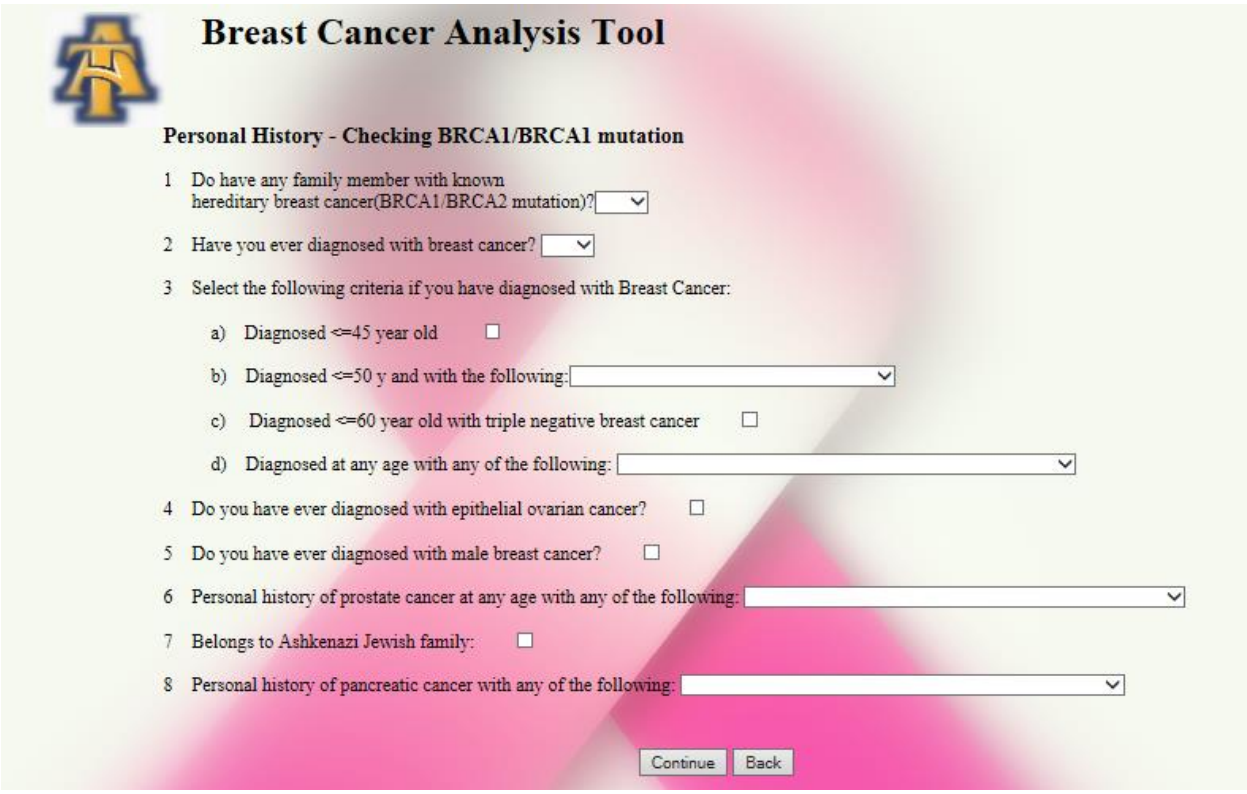

Figure 14. Personal history-BRCA1/BRCA2 analysis.

5.3.1.2.1 Personal history of breast cancer. The personal history of breast cancer has been divided into four different age groups: ages 45 and less, age less than or equal to 50, age less than or equal to 60, and diagnosed at any age. The tool considers the result positive (at risk) if any individual with a personal history of breast cancer meets any of the following criteria.

- Diagnosed at the age of 45 or less: Hereditary breast cancers are likely to appear during early stages of life. Identifying breast cancer before or at the age 45 is an indication of inherited gene mutation.
- Diagnosed at the age of 50 or less along with two breast primaries, one or more close relative with breast cancer at any age, or unknown family history. Primary breast cancer means that the breast cancer is not yet attacked other parts of the body. Women diagnosed

with a primary breast cancer is at higher risk of developing a second primary cancer in another breast.

- Diagnosed at the age of 60 or less with triple-negative breast cancer. Triple-negative breast cancer is a term that use in pathology. In pathology reports the negative results mean that the cancer is not associated with HER2 (HER2-), estrogen receptors (ER-), and progesterone receptors (PR-).
- Diagnosed at any age with any of the following:
    - One or more close relatives with breast cancer diagnosed at the age of 50 years or younger,
    - A relative with epithelial cancer at any age,
    - Two or more relatives with pancreatic cancer or prostate cancer,
    - A close relative with male breast cancer,
    - Ashkenazi ancestry.

Ashkenazi Jewish people are originally from Central or Eastern Europe. They were geographically and genetically isolated people for centuries. Therefore, faulty genes might have been passed down through generations and as a result, put them more at risk. The lifetime risk of getting breast cancer is 90% for any person who belongs to Ashkenazi Jewish ancestry with mutated gene of BRCA. According to NCCN guidelines, if the patient is with personal history of breast cancer and with an Ashkenazi Jewish ancestry, no additional family history is needed to meet the criteria (Daly et al., 2014).

5.3.1.2.2 Personal history of epithelial ovarian cancer. Germline mutations of BRCA1/BRCA2 genes are related to high risk of epithelial ovarian cancer (5% to 10%). A person diagnosed with epithelial cancer at any age will be at risk of hereditary breast cancer.

5.3.1.2.3 Personal history of male breast cancer. Personal history of male breast cancer can be considered a potential factor related to hereditary breast cancer. Men with breast cancer are more susceptible to BRCA2 gene mutation than BRCA1 mutation. The lifetime risk of breast cancer with BRCA2 gene is 6% by age of 70, and it is only 1% with the BRCA1 gene.

5.3.1.2.4 Personal history of pancreatic or prostate cancer. Mutation in BRCA1 or BRCA2 contributes high risk of pancreatic and prostate cancer. Men with BRCA1 or BRCA2 mutations have higher risk of developing prostate cancer. Both men and women with BRCA1 or BRCA2 mutations have higher risk of pancreatic cancer (NCI, 2015). Especially, BRCA2 mutation exhibits higher risk than BRCA1 in prostate cancer with Gleason score greater than eight. Both genes are also associated with high risk of developing pancreatic cancer. As a result, the result will be “at risk” if the person has a personal history of prostate cancer or pancreatic cancer at any age, and has two or more close relatives with breast, ovarian, pancreatic, or prostate cancer. For pancreatic cancer, if the person belongs to Ashkenazi Jewish ancestry, only one additional affected relative is sufficient to make the person at risk (Daly et al., 2014

The result of the personal history analysis will be stored and will continue to check family history. The table below contains different test cases that were created for this section.

Table 3

Test Cases for Personal History-BRCA1/BRCA2 Analysis

Test cases	Do have any family member with known hereditary breast cancer (BRCA1/BRCA2 mutation)?	Have you ever diagnosed with breast cancer?	Personal history of breast cancer and belongs to:				Do you have ever diagnosed with epithelial ovarian cancer?	Do you have ever diagnosed with male breast cancer?	Personal history of prostate cancer with any of the following:	Belongs to Ashkenazi Jewish family	Personal history of pancreatic cancer at any age with any of the following:	Expected result
			<=45 old	<= 50 old	<=60 tripe –ve	at any age with						
P1	Yes	No										At risk
P2	No	Yes	✓									At risk
P3	No	yes		additional primary								At risk
P3	No	Yes		>=1 relatives with BC								At risk
P4	No	Yes		Unknown family history								At risk
P5	No	Yes			✓							At risk
P6	No	Yes				>=1 close relatives with BC diagnosed at the age of <=50 old						At risk

Table 3

Cont.

P7	No	Yes				>=2 relatives with breast cancer						At risk
P8	No	Yes				>=1 close relatives with epithelial OC						At risk
P9	No	Yes				>=2 relatives with pancreatic cancer						At risk
P10	No	Yes				>=2 relatives with prostate cancer						At risk
P11	No	Yes				close relative with male breast cancer						At risk
P12	No	Yes					✓					At risk
P13	No	Yes						✓				At risk
P14	No	Yes							>=2 relatives with breast cancer			At risk
P15	No	Yes							>=2 relatives with pancreatic cancer			At risk
P16	No	Yes							>=2 relatives with prostate cancer			At risk

Table 3

Cont.

P17	No	Yes							>=2 close relatives with epithelial OC			At risk
P18	No	Yes							One more additional relative	✓		At risk
P19	No	Yes								✓		At risk
P20	No	Yes									>=2 relatives with breast cancer	At risk
P21	No	Yes									>=2 relatives with pancreatic cancer	At risk
P22	No	Yes									>=2 relatives with prostate cancer	At risk
P23	No	Yes									>=2 close relatives with epithelial OC	At risk

BC stands for breast cancer, OC stands for epithelial ovarian cancer, PC-pancreatic cancer, Prc-prostate cancer, MBC- Male breast cancer. All of these test cases are for people who are ‘at risk’, and any other conditions or combinations would give the result as ‘not at risk ‘.

5.3.1.3 Family history. In the next section, the program checks the family history. The tool examines the family history and determines the risk based on the frequency of cancer appearing in first-, second-, or third-degree relatives. The output will be stored based on the analysis of the family history. This section has been separated into two subsections to qualify the different conditions for the diagnosis.

5.3.1.3.1 First- or second-degree relatives. First or second-degree relatives who satisfy the conditions. The conditions are same as the criteria used for checking personal history, and it has been discussed earlier in the personal history section.

5.3.1.3.2 Third-degree relatives. Third-degree relatives with breast cancer and or ovarian cancer with two or more close blood relative with breast cancer (one at least before the age of 50) or ovarian cancer.

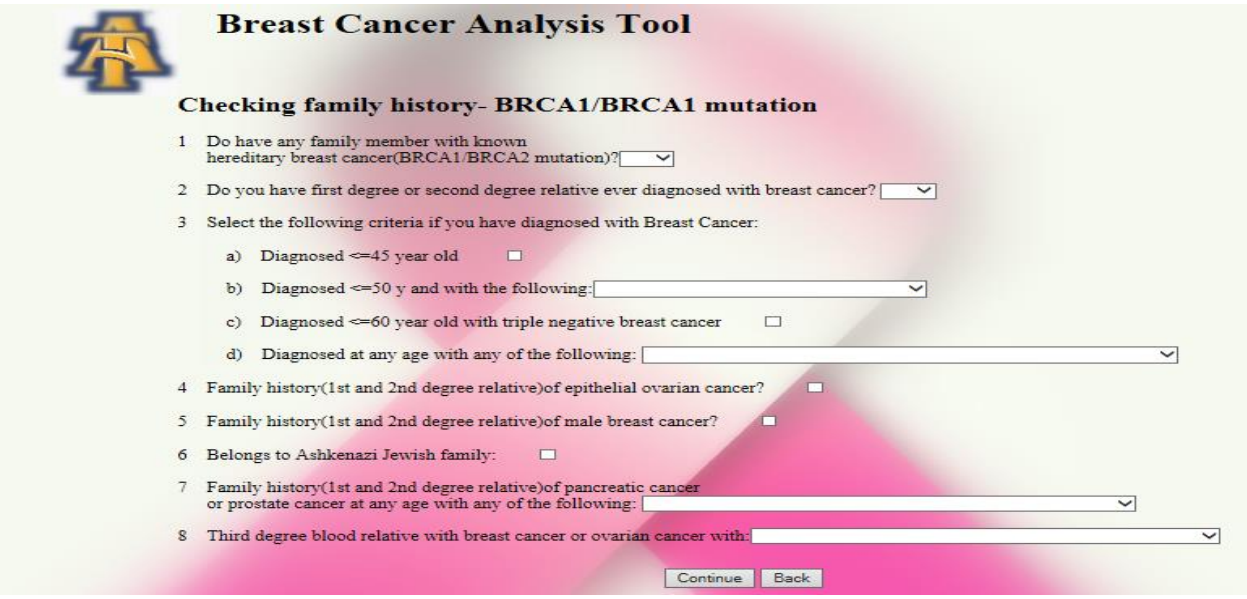

Figure 15. Family history-BRCA1/BRCA analysis.

Table 4

Test Cases for Family History –BRCA1/BRCA2 Analysis

Test cases	Known mutation	1st/2nd degree relatives with breast cancer	the age at diagnosed				Family History				Third degree relative with		Expected Result
			<= 45	<=50	<=60	any age	Epithelial OC	Pancreatic	Prostate	Ashkenazi	Breast cancer	Ovarian cancer	
1	Yes	No											At risk
2	No	Yes	✓	Additional primary									At risk
3	No	Yes		>=1 relatives with BC									At risk
4	No	Yes		Unknown family history									At risk
5	No	Yes			Triple -ve BC								At risk
6	No	Yes				>=1 close relatives with BC diagnosed at the age of <=50 old							At risk
7	No	Yes				>=2 relatives with breast cancer							At risk
8	No	Yes				>=1 close relatives with epithelial OC							At risk

Table 4

Cont.

9	No	Yes				>=2 relatives with pancreatic cancer							At risk
10	No	Yes				>=2 relatives with prostate cancer							At risk
11	No	Yes				close relative with male breast cancer							At risk
12	No	Yes											At risk
13	No	Yes											At risk
14	No	Yes						>=2 relatives with breast cancer					At risk
15	No	Yes						>=2 relatives with pancreatic cancer					At risk
16	No	Yes						>=2 relatives with prostate cancer					At risk
17	No	Yes						>=2 close relatives with epithelial OC					At risk
18	No	Yes						One more additional relative	✓				At risk

Table 4

Cont.

19	No	Yes							✓				At risk
20	No	Yes							>=2 relatives with breast cancer				At risk
21	No	Yes							>=2 relatives with pancreatic cancer				At risk
22	No	Yes							>=2 relatives with prostate cancer				At risk
23	No	Yes							>=2 close relatives with epithelial OC				At risk
24	No	Yes									>=2 relatives with ovarian cancer		At risk
25	No	Yes										>=2 relatives with breast cancer	At risk

The above test case are for family history section. The test cases checks different scenarios for different age groups. In all of the above conditions, the expected output is ‘at risk’. The result will be stored as the program enters the next section.

The risk of breast cancer increases if the person has close blood relatives with a history of breast cancer. The risk increases as the number of affected relatives increases and also varies with people's race and ethnicity. The next section analyzes Li-Fraumeni syndrome (Tp53).

5.3.2 Li-Fraumeni syndrome analysis. At this stage, the software tool checks the presence of Li-Fraumeni syndrome. Breast cancer is common in women with Li-Fraumeni Syndrome (LFS). This syndrome is associated with germline mutation of Tp53 gene. The related tumors and cancers associated with Li-Fraumeni Syndrome are soft-tissue sarcomas, osteosarcomas, premenopausal breast cancer, adrenal cortex, colon cancer, acute leukemia, and brain tumors. The core cancers related with LFS syndrome are sarcoma, adrenocortical carcinoma, breast cancer and brain tumors. Chompret criteria and Classic Li-Fraumeni Syndrome criteria are the two different criteria that have been widely used to check the Tp53 mutation. The flowchart to check Tp53 mutation (Li-Fraumeni syndrome) is given in flowchart section b. The program can be divided into four parts based on different checking criteria. In this section, the software tool checks four different conditions: a) an individual with known mutations, b) classic Li-Fraumeni syndrome criteria, c) Chompret criteria, and d) early age onset criteria. The result will not be stored if the conditions are not met. However, in either condition (criteria met or not), the program will enter into the next section. The analysis has been done as follows:

5.3.2.1 Individual from a family with known mutation. In this part, the program checks if the individual is from a family that is already known to have or is diagnosed with the mutation of Tp53. If the individual has an already identified the mutation of Tp53, then the program would consider the person at risk and enter into the next section after storing the result. According to NCCN guidelines, people who are already diagnosed with Tp53 gene mutations are at risk of

hereditary breast cancer. People who carry the Tp53 mutated gene have a 21% to 49% risk of developing breast cancer by the age of 30, and the lifetime cancer risk is 68% to 93%.

5.3.2.2 Checking classic Li-Fraumeni Syndrome criteria. To meet NCCN criteria for classic Li-Fraumeni Syndrome Criteria, the person should satisfy all three conditions.

- Individual is diagnosed with sarcoma (cancer related to connective tissues) before the age of 45 years; and
- Individual has a first-degree relative with cancer, and it was diagnosed before the age of 45; and
- Individual has one or more first- or second-degree relatives with cancer diagnosed before the age of 45 or with a sarcoma at any age.

5.3.2.3 Checking Chompret criteria. This part the program analyzes three conditions given below. If any of the conditions are met, the output will be stored and the program will exit from the section.

- Individual with a tumor from LFS tumor spectrum (soft-tissue sarcoma, osteosarcoma (cancer develops in bone), brain tumor, breast cancer, adrenocortical carcinoma, leukemia, lung Broncho alveolar cancer) before 46 years of age, and one or more first- or second-degree relatives with cancers that included in the LFS spectrum before the age of 56 years or with multiple primaries at any age; or
- Individual with multiple tumors and two of them belonging to LFS spectrum and diagnosed before the age of 46 years; or
- If the individual was diagnosed with adrenocortical carcinoma (a cancer that begin at the outer layer of the adrenal gland) or choroid plexus carcinoma (tumor arises from brain tissues) at any age of onset, regardless of the family history.

5.3.2.4 Early-age-onset criteria. An individual diagnosed with breast cancer before the age of 35 years can be considered as a person at risk of Tp53 mutation. A person with breast cancer under the age of 30 and with a family history of one or more core cancers in first- or second-degree relative has a 100% chance of having the p53 mutation.

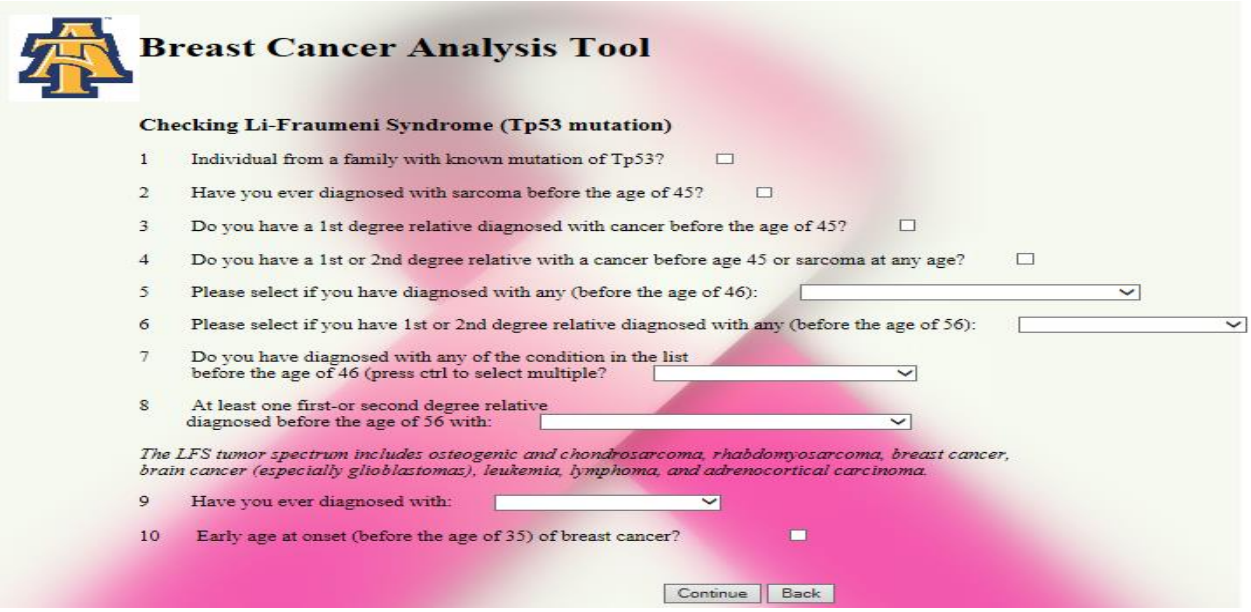

Figure 16. Tp53 analysis.

The above figure is a screenshot from the Tp53 mutation analysis. This section includes ten questionnaires to check the risk of mutated gene. Based on the user's selections the program will check the risk of Tp53 mutation. The accuracy of Chompret criteria alone is 92%. However, the combination of classic Li-Fraumeni syndrome criteria and Chompret criteria exhibits 99% accuracy (Gonzalez et al., 2009).

Table 5

Test Cases for Tp53 Analysis

Test cases	Individual from a known mutation?	Have you ever diagnosed with sarcoma before the age of 45?	Do you have a 1st degree relative diagnosed with cancer before the age of 45?	Do you have a 1st or 2nd degree relative with a cancer before age 45 or sarcoma at any age?	Please select if you have diagnosed with any (before the age of 46):	Please select if you have 1st or 2nd degree relative diagnosed with any (before the age of 56):	Do you have diagnosed with any of the condition in the list before the age of 46?	Have you ever diagnosed with:	Breast cancer diagnosed before the age of 35?	Expected Result *
P1	✓	✓	✓							At risk
P2		✓								Not at risk
P3		✓	✓	✓						At risk
P4					Soft tissue sarcoma	Brain tumor				At risk
P5						Brain tumor				Not at risk
P6					Osteosarcoma	Leukemia				At risk
P7					Breast cancer	Soft tissue sarcoma				At risk
P8					Adrenocortical carcinoma	Leukemia				At risk
P9					Leukemia	Soft tissue sarcoma				At risk
P10					Lung Broncho alveolar cancer	Adreno cortical carcinoma				At risk
P11						Leukemia	Soft Tissue Sarcoma, Brain tumor			At risk

Table 5

Cont.

P12							Multiple breast tumors			Not at risk
P13								Adrenocortical carcinoma		At risk
P14								Choroid plexus carcinoma		At risk
P15									✓	At risk
P16						Leukemia				Not at risk
P17					Soft tissue sarcoma					Not at risk
P18					Osteosarcoma					Not at risk
P19					Breast cancer					Not at risk
P20					Adrenocortical carcinoma					Not at risk
P21					Leukemia					Not at risk
P22					Lung Broncho alveolar cancer					Not at risk

The above table includes different test cases followed for Tp53 analysis section.

5.3.3 Cowden syndrome analysis. In this part, the program checks PTEN (Cowden syndrome) mutation. Cowden syndrome is related to mutations in the PTEN gene. In order to calculate the risk, the flowchart will check four criteria. There are some major and minor criteria that should satisfy the person to be at risk.

Major criteria include:

- Breast cancer
- Endometrial cancer – Cancer begins in the inner layer of the uterus.
- Follicular thyroid cancer
- Multiple GI hamartomas or ganglioneuromas – Cancer starts in the nerve cells
- Macrocephaly (58 cm in adult women, 60 cm in adult men) – A condition in which the head circumference is larger than normal.
- Macular pigmentation of glans penis
- Mucocutaneous lesions with any of the following:
    - One biopsy-proven trichilemmoma (tumor begins in the hair follicle)
    - Multiple palmoplantar keratosis (abnormal thickening of palms and soles)
    - Multifocal or extensive oral mucosal papillomatosis (a skin surface elevation)
    - Multiple cutaneous facial papules (a skin disorder)

Minor criteria include:

- Autism spectrum disorder
- Colon cancer
- Three or more esophageal glycogenic acanthoses (white plaque found in esophagus)
- Lipomas – A tumor which is composed of body fat.
- Mental retardation

- Papillary or Follicular variant of papillary thyroid cancer
- Thyroid structural lesions
- Renal cell carcinoma – A type of kidney cancer
- Single GI hamartoma or ganglioneuroma
- Testicular lipomatosis
- Vascular anomalies

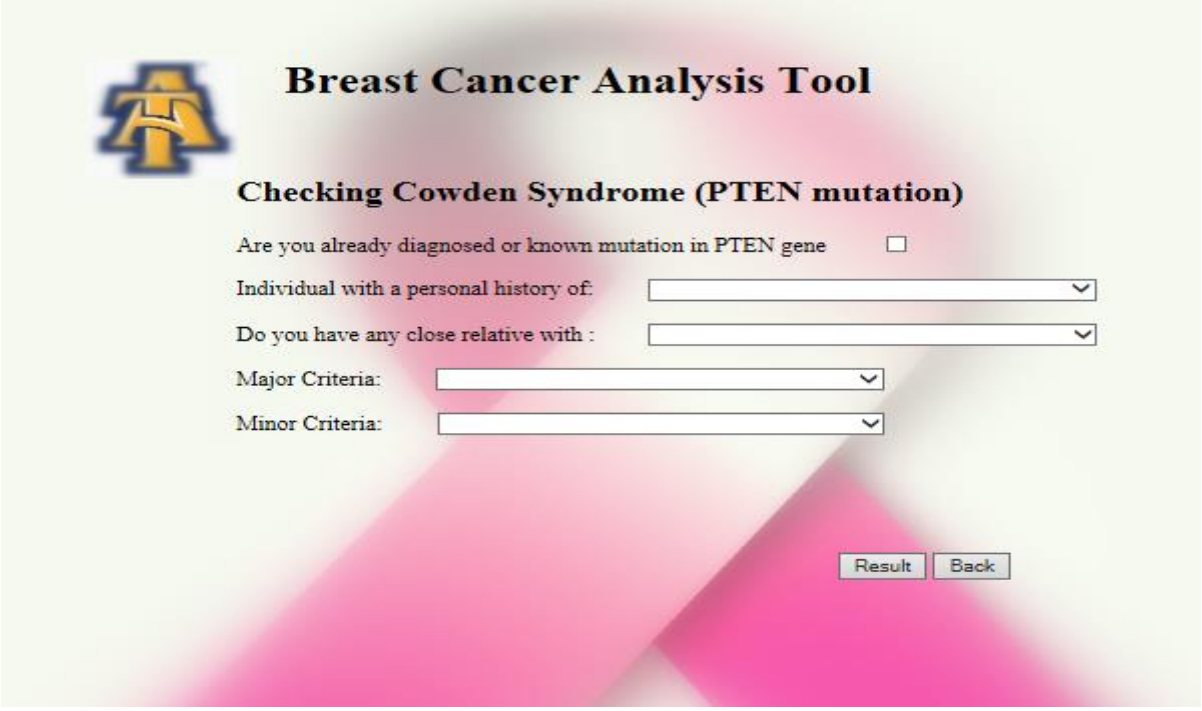

Figure 17. PTEN analysis.

The above figure is a screenshot of PTEN analysis section. As shown in the figure, the program checks individual known mutation history, personal history and family history. The major and minor criteria dropdowns include different clinical criteria for PTEN gene testing. The program determines the risk of mutation based on the selections that have been made by the user.

Table 6

Test Cases of PTEN Analysis

Test cases	Are you already diagnosed or known mutation in PTEN gene	Individual with a personal history of:	Do you have any close relative with:	Major Criteria:	Minor Criteria:	Expected Result
P1	✓					At risk
P2		Bannayan-Riley-Ruvalcaba syndrome				At risk
P3		Adult Lhermitte-Ducolos diseases				At risk
P4		Autism spectrum disorder and macrocepahly				At risk
P5		Two or more biopsy-proven trichilemmoas				At risk
P6				Breast cancer Endometrial cancer		At risk
P7				Follicular thyroid cancer Multiple GI hamartomas or gangalloneuromas		At risk
P8				Macrocephaly Macular pigmentation of glans penis		At risk
P9				Mucocutaneous lesions and One biopsy-proven trichilemmoma		At risk

Table 6.

Cont.

P10				Mucocutaneous lesions and Multiple palmoplantar keratosis		At risk
P11				Mucocutaneous lesions and Multiple cutaneous facial papules		At risk
P12				Breast cancer	Autism spectrum disorder, Colon cancer, Three or more esophageal glycogenic acanthoses, Lipomas	At risk
P13				Endometrial cancer	Lipomas, Mental retardation, Colon cancer	At risk
P14				Follicular thyroid cancer	Lipomas, Mental retardation, Colon cancer	At risk
P15					Autism spectrum disorder, Colon cancer, Lipomas, Mental retardation, Papillary or Follicular variant of papillary thyroid cancer	At risk
P16			Breast cancer Endometrial cancer			At risk
P17			Multiple GI hamartomas or gangalioneuromas Macrocephaly			At risk

The above table-6 include different test criteria for PTEN gene mutations.

The program has four sections: 1) individual from a family with an already known mutation, 2) individual meeting clinical diagnostic criteria, 3) individual with personal history and other NCCN criteria 4) individual with a close blood relative meeting NCCN criteria.

5.3.3.1 Individual with an already known mutation. First the flowchart will check if the individual has a known mutation of PTEN gene.

5.3.3.2 Personal history. Individual with a personal history of any of the following diseases:

- Bannayan-Riley-Ruvalcaba syndrome (BRRS)
- Adult Lhermitte-Ducolos diseases
- Autism spectrum disorder and macrocepahly
- Two or more biopsy-proven trichilemmoas
- Two or more major criteria
- Three major criteria, without macrocepahly
- One major and more than three minor criteria
- More than four minor criteria

5.3.3.3 Family history. The program check the family history of close blood relatives with any one or more major criteria or two or more minor criteria.

The test result will be stored before it enters into the next section.

5.3.4 Forecasting cancer in future offspring. Every child inherits two copy from their parents; one from their mother and one from the father. BRCA1/BRCA2, Tp53 and PTEN genes are not associated with x-linked inheritance. Therefore, the mutations can be inherited from either parent. Offspring of the individual with any of these gene mutations can inherit the mutation.

As BRCA1, BRCA2, Tp53 and PTEN genes exhibit autosomal inheritance pattern, there are four possible combinations for the children's gene copy. As shown in figure 3, two of the four, or 50 percent, can inherit the mutated genes (aA, aA). The other 50 percent do not inherit the mutated genes (AA, AA). These four combinations are possible every time a pregnancy occurs. The gender of the children (whether they are sons or daughters) does not matter. The chance of getting normal child is 50/50 for each pregnancy. Consider a mother who carries BRCA1 mutated gene, and her partner carries normal genes. In each pregnancy, the chances of getting normal child is 50% and also chances of getting child with the abnormal gene is also 50%. We are using this information to predict the chances of hereditary breast cancer in the offspring or the child to be born if the individual is analyzed at risk by this tool.

5.3.5 Output. In the final session, the output from different sections will be combined and displayed along with the next generation's risk. There will be different output based on the different sections. The result contains mutated gene risk and breast cancer risk. The result also displays the risk of passing mutated genes to next generation. In the output, it has been advised that the result is not a substitution for professional genetic test. If the individual is at risk, we recommend a discussion with an oncologist for better treatment. This program will have four different outputs as shown in below. The outputs are based on different gene mutations analysis. It includes BRCA1/BRCA2 analysis, Tp53 analysis, PTEN analysis and 'not at risk' situations. The user can print the result and consult an oncologist for further assessments and treatments.

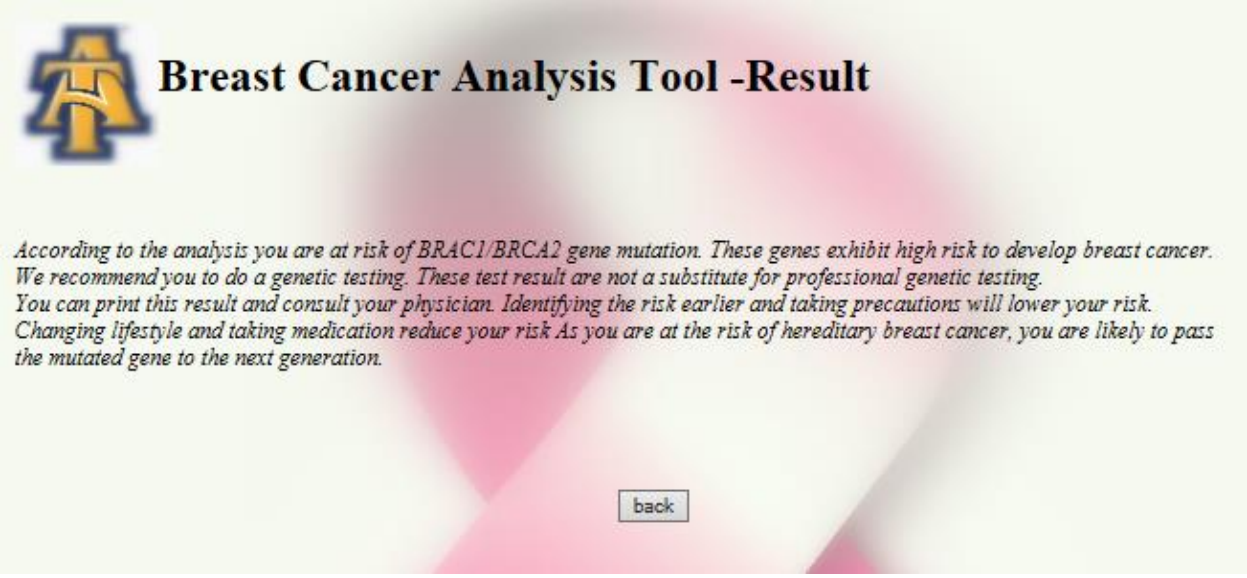

Figure 18. Output – BRCA1/BRCA2 mutation.

The above figure represents one of the output. In this scenario, the person has met family history criteria of HBOC syndrome analysis.

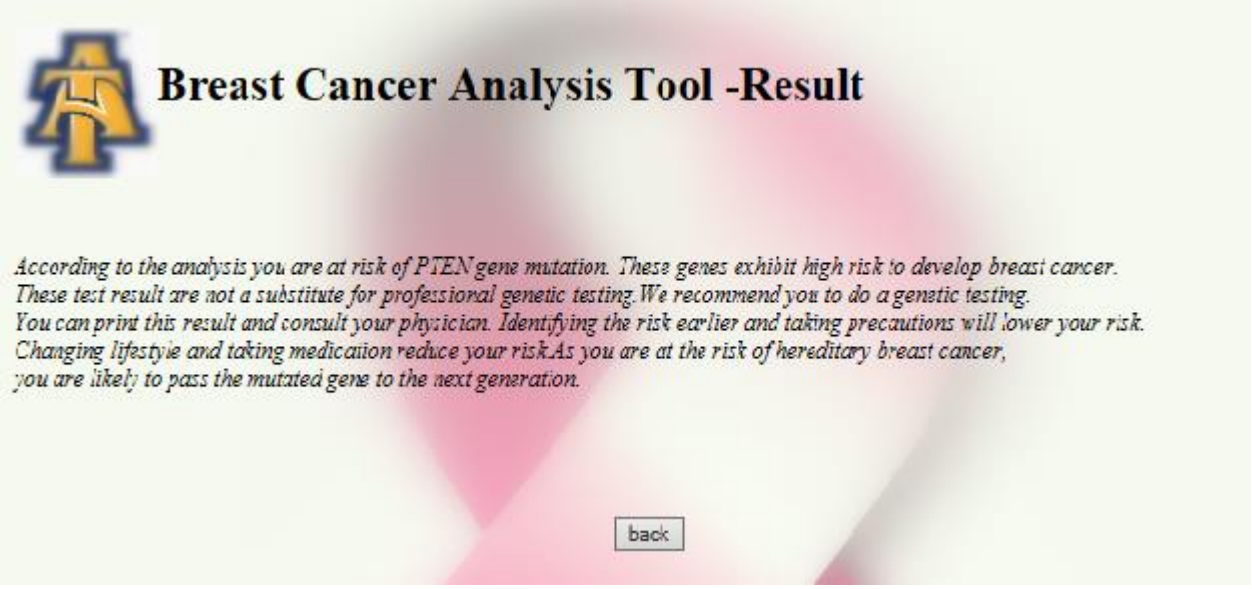

Figure 19. Output – PTEN mutation.

The above figure shows the analysis result from Cowden syndrome or PTEN gene mutation.

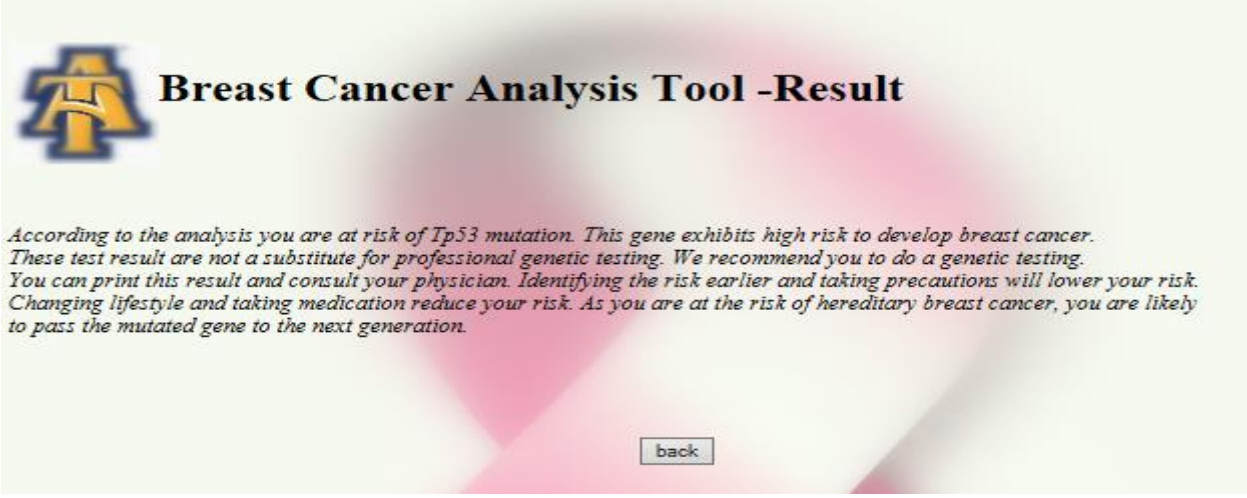

Figure 20. Output – Tp53 mutation.

The figure- 15 is a screen image of the output of the program for a user who met the Tp53 mutation criteria.

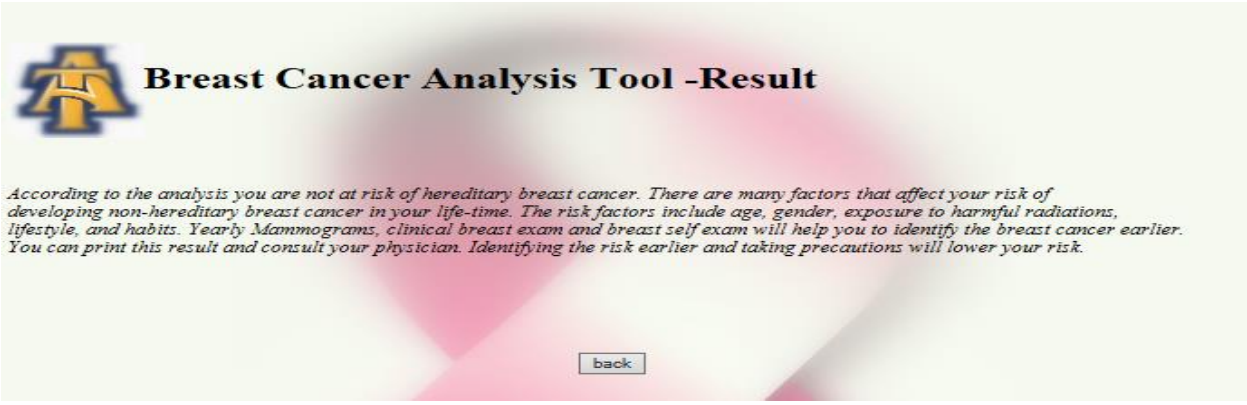

Figure 21. Output – 'not at risk' case.

The above figure represents the output for a user who does not meet any criteria.

5.3.6 Features of the program.

- Includes PTEN and Tp53 genes: This tool includes high-risk Tp53 and PTEN genes, and none of the existing breast cancer tools available consider these high risk genes.
- Consider ovarian, prostate, pancreatic cancer history.

- Included Ashkenazi Jewish history.
- Available to the public: Many of the existing tools are proprietary and are not available to the public. This tool is open source, and it will be available to the public.
- User-friendly: This is a very easy-to-use software tool, and any person with basic knowledge in computer operation can easily use it.
- Easy to understand: It is easily understood by people who have little breast cancer awareness.
- Universal acceptance: This tool is designed based on NCCN guidelines, and it is not restricted to a set of people based on their demography or ethnic group. The guidelines have been developed in coordination with twenty-five global leading cancer centers. Most of the existing tools have been developed based on the data collected from a particular demography.
- Includes third-degree relatives: Considers family history that includes first-, second-, and third-degree relatives. Many tools available are not considering third-degree relatives and family history.

Early detection and screening would help to prevent the cancer or to find the disease at an early age. Taking drugs would prevent or delay cancer onset if you are at risk of hereditary breast cancer syndrome. For example, medicines such as Tamoxifen and Raloxifene lower the breast cancer risk in the general population (Moyer, 2013). Individuals at risk may choose to have surgery to reduce their risk. The individual at risk is strongly recommended to have breast or ovarian cancer screening, including mammograms, MRI exams, ultrasound, and breast exams, performed by a physician.

CHAPTER 6

Future Work and Conclusion

Predicting disease with computer modeling and mathematical analysis is becoming more popular. Combining all known diseases and diagnosis would help medical practitioners to determine and predict diseases in the current or next generations. As the years go by, family health history stored in computers provides better assessments, accuracy, and standardization in forecasting diseases. Computers could be trained to forecast and predict specific diseases. We want to integrate this tool with other common inherited diseases such as cystic fibrosis, Down syndrome, type 2 diabetes, inherited heart diseases, or other hereditary cancers, and that will help both common people and medical professionals. We want to build this prototype as a comprehensive computer application to predict most of the common inheritable and predictable diseases. We like to add more hereditary cancers and improve the existing program with other moderate risk genes such as PALB1, BRIP, and LKB1. We want to include the non-hereditary risk factors that are related to lifestyle and diet. We also like to work on all known hereditary and non-hereditary cancers and integrate them with this tool. In future, we are considering to integrate this application with smartphone applications.

References

Abeliovich, D., Kaduri, L., Lerer, I., Weinberg, N., Amir, G., Sagi, M., ... & Peretz, T. (1997). The founder mutations 185delAG and 5382insC in BRCA1 and 6174delT in BRCA2 appear in 60% of ovarian cancer and 30% of early-onset breast cancer patients among Ashkenazi women. American journal of human genetics, 60(3), 505.

Agrawal, R., Suleiman, M., Seay, C., & Gloster, C. (2013, September). Dynamic Disease Forecast Network Using Family Medical History. In Healthcare Informatics (ICHI), 2013 IEEE International Conference on (pp. 486-486). IEEE.

Agalliu, I., Gern, R., Leanza, S., & Burk, R. D. (2009). Associations of high-grade prostate cancer with BRCA1 and BRCA2 founder mutations. Clinical Cancer Research, 15(3), 1112-1120.

Aiello-Laws, L. (2011, February). Genetic cancer risk assessment. In *Seminars in oncology nursing* (Vol. 27, No. 1, pp. 13-20). WB Saunders.

Allain, D. C. (2008). Genetic counseling and testing for common hereditary breast cancer syndromes: a paper from the 2007 William Beaumont hospital symposium on molecular pathology. *The Journal of Molecular Diagnostics*, *10*(5), 383-395.

American Cancer Society. (2014). What is breast cancer. Retrieved from http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-what-is-breast-cancer

American Cancer Society. (2014). Family Cancer Syndromes. Retrieved from http://www.cancer.org/cancer/cancercauses/geneticsandcancer/heredity-and-cancer

American Cancer Society. (2014). What are the risk factors for breast cancer?. Retrived from, http://www.cancer.org/cancer/breastcancer/

American Cancer Society. (2015). Cancer Facts and Figures 2015. Retrived from http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf

Amir, E., Evans, D. G., Shenton, A., Lalloo, F., Moran, A., Boggis, C., ... & Howell, A. (2003). Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. *Journal of medical genetics*, *40*(11), 807-814.

Antoniou, A. C., Cunningham, A. P., Peto, J., Evans, D. G., Lalloo, F., Narod, S. A., … Easton, D. F. (2008). The BOADICEA model of genetic susceptibility to breast and ovarian cancers:

Antoniou, A., Pharoah, P. D. P., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J. L., ... & Easton, D. F. (2003). Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *The American Journal of Human Genetics*, *72*(5), 1117-1130.

Atchley, D. P., Albarracin, C. T., Lopez, A., Valero, V., Amos, C. I., Gonzalez-Angulo, A. M., ... & Arun, B. K. (2008). Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer. Journal of Clinical Oncology, 26(26), 4282-4288.

Breastcancer.org. 2014. Retrieved from http://www.breastcancer.org/risk/factors/family_history

Biswas, S., Tankhiwale, N., Blackford, A., Barrera, A. M. G., Ready, K., Lu, K., & Arun, B. (2012). Assessing the added value of breast tumor markers in genetic risk prediction model BRCAPRO. *Breast cancer research and treatment*, *133*(1), 347-355

Boughey, J. C., Hartmann, L. C., Anderson, S. S., Degnim, A. C., Vierkant, R. A., Reynolds, C. A., … Pankratz, V. S. (2010). Evaluation of the Tyrer-Cuzick (International Breast

Cancer Intervention Study) Model for Breast Cancer Risk Prediction in Women With Atypical Hyperplasia. *Journal of Clinical Oncology, 28*(22), 3591–3596. doi:10.1200/JCO.2010.28.0784

Cancer Research U.K. (2014) Breast cancer genes. Retrieved from http://www.cancerresearchuk.org/about-cancer/type/breast-cancer/about/risks/breast-cancer-genes.

Centers for Diseases Control and Prevention. (2014). What Are the Risk Factors for Breast Cancer?. Retrieved from http://www.cdc.gov/cancer/breast/basic_info/risk_factors.htm

Chapman, D. D. (2007, February). Cancer genetics. In *Seminars in oncology nursing* (Vol. 23, No. 1, pp. 2-9). WB Saunders.

Chen, S., & Parmigiani, G. (2007). Meta-analysis of BRCA1 and BRCA2 penetrance. Journal of Clinical Oncology, 25(11), 1329-1333.

Claus, E. B., Risch, N., & Thompson, W. D. (1993). The calculation of breast cancer risk for women with a first degree family history of ovarian cancer. Breast cancer research and treatment, 28(2), 115-120.

Claus, E. B., Risch, N., & Thompson, W. D. (1991). Genetic analysis of breast cancer in the cancer and steroid hormone study. *American journal of human genetics*, *48*(2), 232.

Comen, E., Davids, M., Kirchhoff, T., Hudis, C., Offit, K., & Robson, M. (2011). Relative contributions of BRCA1 and BRCA2 mutations to "triple-negative" breast cancer in Ashkenazi women. Breast cancer research and treatment, 129(1), 185-190

Daly, M. B., Pilarski, R., Axilbund, J. E., Buys, S. S., Crawford, B., Friedman, S., ... & Kumar, R. (2014). Genetic/Familial high-risk assessment: breast and ovarian, version 1.2014. *Journal of the National Comprehensive Cancer Network*, *12*(9), 1326-1338.

Ellsworth, R. E., Decewicz, D. J., Shriver, C. D., & Ellsworth, D. L. (2010). Breast cancer in the personal genomics era. *Current genomics*, *11*(3), 146.

Eng, C. (2014). PTEN hamartoma tumor syndrome (PHTS).

Eng, C. (2000). Will the real Cowden syndrome please stand up: revised diagnostic criteria. *Journal of medical genetics*, *37*(11), 828-830.

Evans, D. G., & Howell, A. (2007). Breast cancer risk-assessment models. *Breast Cancer Res*, *9*(5), 213.

Evans, D. G. R., Susnerwala, I., Dawson, J., Woodward, E., Maher, E. R., & Lalloo, F. (2010). Risk of breast cancer in male BRCA2 carriers. *Journal of medical genetics*, *47*(10), 710-711.

Fostira, F., Tsitlaidou, M., Papadimitriou, C., Pertesi, M., Timotheadou, E., Stavropoulou, A. V., ... & Fountzilas, G. (2012). Prevalence of BRCA1 mutations among 403 women with triple-negative breast cancer: implications for genetic screening selection criteria: a Hellenic Cooperative Oncology Group Study. Breast cancer research and treatment, 134(1), 353-362.

Frank, T. S., Manley, S. A., Olopade, O. I., Cummings, S., Garber, J. E., Bernhardt, B., ... & Thomas, A. (1998). Sequence analysis of BRCA1 and BRCA2: correlation of mutations with family history and ovarian cancer risk. *Journal of Clinical Oncology*, *16*(7), 2417-2425.

Gail, M. H., Costantino, J. P., Bryant, J., Croyle, R., Freedman, L., Helzlsouer, K., & Vogel, V. (1999). Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *Journal of the National Cancer Institute*, *91*(21), 1829-1846.

Gail, M. H., Brinton, L. A., Byar, D. P., Corle, D. K., Green, S. B., Schairer, C., & Mulvihill, J. J. (1989). Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. Journal of the National Cancer Institute, 81(24), 1879-1886.

Gonzalez-Angulo, A. M., Timms, K. M., Liu, S., Chen, H., Litton, J. K., Potter, J., ... & Meric-Bernstam, F. (2011). Incidence and outcome of BRCA mutations in unselected patients with triple receptor-negative breast cancer. Clinical Cancer Research, 17(5), 1082-1089

Gonzalez, K. D., Noltner, K. A., Buzin, C. H., Gu, D., Wen-Fong, C. Y., Nguyen, V. Q., ... & Weitzel, J. N. (2009). Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations. Journal of Clinical Oncology, 27(8), 1250-1256.

Gonzalez, K. D., Buzin, C. H., Weitzel, J. N., & Sommer, S. S. (2009). Reply to J. Tinat et al. Journal of Clinical Oncology, 27(26), e110-e110.

Huo, D., Senie, R. T., Daly, M., Buys, S. S., Cummings, S., Ogutha, J., ... & Olopade, O. I. (2009). Prediction of BRCA mutations using the BRCAPRO model in clinic-based African American, Hispanic, and other minority families in the United States. Journal of Clinical Oncology, 27(8), 1184-1190.

Howe, H. L., Weinstein, R., Alvi, R., Kohler, B., & Ellison, J. H. (2005). Women with multiple primary breast cancers diagnosed within a five year period, 1994–1998. Breast cancer research and treatment, 90(3), 223-232.

Jazaeri, A. A., Lu, K., Schmandt, R., Harris, C. P., Rao, P. H., Sotiriou, C., ... & Liu, E. T. (2003). Molecular determinants of tumor differentiation in papillary serous ovarian carcinoma. Molecular carcinogenesis, 36(2), 53-59

Lalloo, F., & Evans, D. G. (2012). Familial breast cancer. *Clinical genetics*, *82*(2), 105-114.

King, M. C., Marks, J. H., & Mandell, J. B. (2003). Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. Science, 302(5645), 643-646.

Kirchhoff, T., Kauff, N. D., Mitra, N., Nafa, K., Huang, H., Palmer, C., ... & Offit, K. (2004). BRCA mutations and risk of prostate cancer in Ashkenazi Jews. Clinical Cancer Research, 10(9), 2918-2921

Lee, A. J., Cunningham, A. P., Kuchenbaecker, K. B., Mavaddat, N., Easton, D. F., & Antoniou, A. C. (2014). BOADICEA breast cancer risk prediction model: updates to cancer incidences, tumour pathology and web interface. *Br J Cancer, 110*(2), 535-545. doi: 10.1038/bjc.2013.730

Lee, L. J., Alexander, B., Schnitt, S. J., Comander, A., Gallagher, B., Garber, J. E., & Tung, N. (2011). Clinical outcome of triple negative breast cancer in BRCA1 mutation carriers and noncarriers. Cancer, 117(14), 3093-3100.

Moyer, V. A. (2013). Medications for risk reduction of primary breast cancer in women: US Preventive Services Task Force recommendation statement. *Annals of internal medicine*, *159*(10), 698-708.

Murray, A. J., & Davies, D. M. (2013). The genetics of breast cancer. *Surgery (Oxford), 31*(1), 1-3.

National Cancer Institute.(2015). Genetics of Breast and Gynecologic Cancers. Retrieved from http://www.cancer.gov/cancertopics/pdq/genetics/breast-and-ovarian/HealthProfessional/page1

National Cancer Institute. (2015). What Is Cancer?. Retrieved from http://www.cancer.gov/cancertopics/what-is-cancer

National Library of Medicine. (2015). What are the different ways in which a genetic condition can be inherited?. Retrived from http://ghr.nlm.nih.gov/handbook/inheritance/inheritancepatterns

National Center for Biotechnology Information. NCBI. Genetic Alliance; The New York-Mid-Atlantic Consortium for Genetic and Newborn Screening Services. Understanding Genetics: A New York, Mid-Atlantic Guide for Patients and Health Professionals. Washington (DC): Genetic Alliance; 2009 Jul 8. CHAPTER 5, GENETIC COUNSELING. Available from: http://www.ncbi.nlm.nih.gov/books/NBK115552/

Palmer, J. R., Hatch, E. E., Rosenberg, C. L., Hartge, P., Kaufman, R. H., Titus-Ernstoff, L., ... & Hoover, R. N. (2002). Risk of breast cancer in women exposed to diethylstilbestrol in utero: preliminary results (United States). *Cancer Causes & Control*, *13*(8), 753-758.

Pagon, R. A., Adam, M. P., Bird, T. D., Dolan, C. R., Fong, C. T., Smith, R. J., ... & Feldman, G. L. (2013). BRCA1 and BRCA2 Hereditary Breast and Ovarian Cancer

Petrucelli, N., Daly, M. B., & Feldman, G. L. (2010). Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. Genetics in Medicine, 12(5), 245-259.

Pilarski, R. (2009). Cowden syndrome: a critical review of the clinical literature. *Journal of genetic counseling*, *18*(1), 13-27.

Pilarski, R., Stephens, J. A., Noss, R., Fisher, J. L., & Prior, T. W. (2011). Predicting PTEN mutations: an evaluation of Cowden syndrome and Bannayan–Riley–Ruvalcaba syndrome clinical features. Journal of medical genetics, jmg-2011.

Riley, B. D., Culver, J. O., Skrzynia, C., Senter, L. A., Peters, J. A., Costalas, J. W., ... & Trepanier, A. M. (2012). Essential elements of genetic cancer risk assessment,

counseling, and testing: updated recommendations of the National Society of Genetic Counselors. *Journal of genetic counseling*, *21*(2), 151-161.

Santaguida, P., Carroll, J., Allanson, J., Culebro, C. R., Brouwers, M., & Raina, P. (2007). *Collection and use of cancer family history in primary care*. US Department of Health and Human Services, Public Health Service, Agency for Healthcare Research and Quality.

Schaffer, J. V., Kamino, H., Witkiewicz, A., McNiff, J. M., & Orlow, S. J. (2006). Mucocutaneous neuromas: an underrecognized manifestation of PTEN hamartoma-tumor syndrome. *Archives of dermatology*, *142*(5), 625-632.

Tan, M. H., Mester, J. L., Ngeow, J., Rybicki, L. A., Orloff, M. S., & Eng, C. (2012). Lifetime cancer risks in individuals with germline PTEN mutations. Clinical Cancer Research, 18(2), 400-407.

Tai, Y. C., Chen, S., Parmigiani, G., & Klein, A. P. (2008). Incorporating tumor immunohistochemical markers in BRCA1 and BRCA2 carrier prediction. *Breast Cancer Res*, *10*(2), 401.

Thorne, H., Willems, A. J., Niedermayr, E., Hoh, I. M., Li, J., Clouston, D., ... & Bolton, D. (2011). Decreased prostate cancer-specific survival of men with BRCA2 mutations from multiple breast cancer families. Cancer Prevention Research, 4(7), 1002-1010.

U.S. Department of Health & Human Services. (2014). My Family Health Portrait. Retrieved from https://familyhistory.hhs.gov/FHH/html/index.html

USA. Department of Health and Human Services.What are the different ways in which a genetic condition can be inherited?. [Online]. Available from,

http://ghr.nlm.nih.gov/handbook/inheritance/inheritancepatterns [Accessed on December 2013]

Williams, J. K. (1997, May). Principles of genetics and cancer. In *Seminars in oncology nursing* (Vol. 13, No. 2, pp. 68-73). WB Saunders.

van Harssel, JJT, van Roozendaal, CEP, Detisch, Y, Brandao, RD, Paulussen, ADC, Zeegers, M, . . . García, EB Gómez. (2010). Efficiency of BRCAPRO and Myriad II mutation probability thresholds versus cancer history criteria alone for BRCA1/2 mutation detection. *Familial cancer*, *9*(2), 193-201.

Yamada, K. M., & Araki, M. (2001). Tumor suppressor PTEN: modulator of cell signaling, growth, migration and apoptosis. Journal of cell science, 114(13), 2375-2382

Yoon, P. W., Scheuner, M. T., Peterson-Oehlke, K. L., Gwinn, M., Faucett, A., & Khoury, M. J. (2002). Can family history be used as a tool for public health and preventive medicine?. *Genetics in Medicine*, *4*(4), 304-310

Yoon, P. W., Scheuner, M. T., & Khoury, M. J. (2003). Research priorities for evaluating family history in the prevention of common chronic diseases. *American journal of preventive medicine*, *24*(2), 128-135.

Young, S. R., Pilarski, R. T., Donenberg, T., Shapiro, C., Hammond, L. S., Miller, J., ... & Narod, S. A. (2009). The prevalence of BRCA1 mutations among young women with triple-negative breast cancer. *BMC cancer*, *9*(1), 86.