CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Bioinformatics and Experimental Analysis of the Genetic and Non-genetic Basis of Breast Cancer in Pakistani Population

by

Raisa Bano

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the

Faculty of Health and Life Sciences
Department of Biosciences

2019

Bioinformatics and Experimental Analysis of the Genetic and Non-genetic Basis of Breast Cancer in Pakistani Population

By

Raisa Bano

(PI091002)

Dr. Omar Bagasra (Professor)

Claflin University, Orangeburg, South Carolina, USA

Dr. Manola KN (Research Scientist)

National Centre for Scientific Research, Aghia Paraskevi, Greece

Prof. Dr. Hamid Rashid

(Thesis Supervisor)

Dr. Sahar Fazal

(Head, Department of Biosciences)

Prof. Dr. Muhammad Abdul Qadir

(Dean, Faculty of Health and Life Sciences)

DEPARTMENT OF BIOSCIENCES

CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

ISLAMABAD

2019

This dissertation is dedicated to my loving husband, Dr. Mohammad Haroon Khan, who encouraged me, and put his academic profession on hold so I could achieve my dream. His good examples have taught me to work hard for the things that I aspire to achieve. Thank you, Haroon, for your love, wisdom and support. To my daughters, Haya Haroon, Safia Haroon and Hania Haroon, not a day did you complain about how busy I was. Also to my parents, my late grandfather, Eng. Akhter Ali and my in-laws, specially my Father-in-law, Mr. Mohammad Moosa Khan, your prayers have been answered. ALLAH has done it again. I give my deepest expression of love and appreciation for the encouragement that you all gave and the sacrifices you made during this Ph.D program. Thank you for your support.

CAPITAL UNIVERSITY OF SCIENCE & TECHNOLOGY
ISLAMABAD

Expressway, Kahuta Road, Zone-V, Islamabad
Phone:+92-51-111-555-666 Fax: +92-51-4486705
Email: info@cust.edu.pk Website: https://www.cust.edu.pk

CERTIFICATE OF APPROVAL

This is to certify that the research work presented in the thesis, entitled "**Bioinformatics and Experimental Analysis of the Genetic and Non-genetic Basis of Breast Cancer in Pakistani Population**" was conducted under the supervision of **Dr. Hamid Rashid**. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the **Department of Biosciences, Capital University of Science and Technology** in partial fulfillment of the requirements for the degree of Doctor in Philosophy in the field of **Biosciences.** The open defence of the thesis was conducted on **28 December, 2018.**

Student Name : Ms. Raisa Bano (PI091002)

Raisa Bano 28/12/18

The Examination Committee unanimously agrees to award PhD degree in the mentioned field.

Examination Committee :

(a) External Examiner 1: Dr. Hussnain Ahmed Janjua
Associate Professor
ASAB, NUST, Islamabad

(b) External Examiner 2: Dr. Amir Ali Abbasi
Associate Professor
NCB, QAU, Islamabad

(c) Internal Examiner : Dr. Sahar Fazal
Associate Professor
CUST, Islamabad

Supervisor Name : Dr. Hamid Rashid
Professor
CUST, Islamabad

Name of HoD : Dr. Sahar Fazal
Associate Professor
CUST, Islamabad

Name of Dean : Dr. Muhammad Abdul Qadir
Professor
CUST, Islamabad

AUTHOR'S DECLARATION

I, **Ms. Raisa Bano (Registration No. PI091002)**, hereby state that my PhD thesis titled, '**Bioinformatics and Experimental Analysis of the Genetic and Non-genetic Basis of Breast Cancer in Pakistani Population**' is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/ world.

At any time, if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my PhD Degree.

Raisa Bano
28/12/18

(Ms. Raisa Bano)

Dated: 28 December, 2018

Registration No : PI091002

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled "**Bioinformatics and Experimental Analysis of the Genetic and Non-genetic Basis of Breast Cancer in Pakistani Population**" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/ cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD Degree, the University reserves the right to withdraw/ revoke my PhD degree and that HEC and the University have the right to publish my name on the HEC/ University Website on which names of students are placed who submitted plagiarized thesis.

Raisa Bano
28/12/18

(**Ms. Raisa Bano**)
Registration No. PI091002

Dated: 28 December, 2018

List of Publications

It is certified that following publication(s) have been made out of the research work that has been carried out for this thesis:-

Journal Articles

1. **Bano R**, Ismail M, Nadeem A, Khan MH, Rashid H. 2016. Potential Risk Factors for Breast Cancer in Pakistani Women. *Asian Pac J Cancer Prev.*, 17 (9): 4307-4312.

Conference Proceedings

1. **Bano R**, Ismail M and Rashid H. (2012). BRCA1 mutations in Pakistani population associated with breast cancer. *National Science Conference-Road Map of Cutting Edge Technologies. PMAS-Arid Agriculture University*, Rwp. Jan 10-12. p. 217.

2. **Bano R**, and Rashid H. (2011). Diversity of BRCA1 Mutations in Breast Cancer in Pakistan. *10^{th} Shaukat Khanum Memorial Cancer Symposium.* November 18-20, Lahore-Pakistan.

Abstracts

1. **Bano R**, Mansoor Q, Rashid H, Ismail M. 2012. Genetic analysis of BRCA1 gene in human male and female breast cancer. *Pak J Physiol.*, 8(Suppl 1): 30.

Raisa Bano
(PI091002)

Acknowledgements

I bow in deep reverence to "Almighty ALLAH", the most merciful, most beneficent, the most adorable, deity and paramount to whom I am indebted. Allah who is omnipresent, omnipotent and omniscient and I bow my head before, "Almighty Allah" for paying my thanks.

A dissertation is not the outcome of the efforts of entirely one individual. Many people have contributed to its development. At this time, I take the opportunity to acknowledge those who have made some impact on my doctoral journey and accomplishment.

I would like to express my gratitude to my dissertation supervisor Prof. Dr. Hamid Rashid for his guidance, instruction, constructive criticism and great patience throughout my PhD study and preparation of this thesis. I am also thankful to my Co-supervisor Dr. Aamer Nadeem for his guidance and support throughout the whole course. I am especially thankful to my foreign collaborators Dr. Jean Baptise-Cazier (Chair Bioinformatics/Director, Centre for Computational Biology, University of Birmingham, UK), Dr. Anastasia Samsonova (Bioinformatics Core Leader, University of Oxford, UK) and Dr. Christiana Kartsonaki (Senior Statistician at the Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, UK), for their kind support and encouragement in the study when I was in University of Oxford under International Research Support Initiative Program (IRSIP) of HEC Pakistan. They really polished my skills and I was exposed to the world top class facilities which broaden my horizon. Without their support and help, this thesis would not have been possible.

I would like to thank Dr. Muhammad Ismail, Mr. Qaisar Mansoor and their team, our collaborator in Institute of Biomedical and Genetic Engineering, Islamabad, Pakistan. I sincerely thank them for their selfless help throughout my PhD study and sharing of ideas with me. They have taught and advised me on various laboratory techniques, their advice and suggestions were invaluable to me.

I would like to acknowledge the scientists and staff members of NORI hospital, Islamabad, Pakistan for allowing me to the hospital facility for sample collection. I would also thank Mr. Haider and Mr. Zahoor of NORI for their help in the collection of blood samples.

Special thanks are also given to all the faculty members of the Department of Biosciences, Capital University of Science and Technology, Islamabad for their encouragement and cooperation.

To all my friends, especially Mrs. Afifa Navid and Mrs. Fauzia Jabeen who have helped me stay sane through these difficult years. Thank you for your understanding and encouragement in my many, many moments of crisis. Your friendship makes my life a wonderful experience. I cannot list all the names here, but you are always on my mind.

I would also like to thank my parents, my brothers specially Mr. Shahid Ali who always supported me and encouraged me with their best wishes. My deepest thanks and gratitude are also extended to my uncles specially Mr. Qaiser Ahmed who stood behind me and supported me every time when needed. Thank you for giving me strength to reach for the stars and chase my dreams. Their love, patience and sacrifice are vested in every page of this study.

I would like to acknowledge the support from Higher Education Commission (HEC) Pakistan for funding and supporting my Ph.D studies through offering me Indigenous Fellowship.

A special group of people, my in-laws are not mentioned yet, because they deserve their own part. I praise the enormous amount of help and support by them throughout my family life. Whenever I felt tired and sense of despair gripped me, I always found them standing by my side, providing me words of immense consolation.

A very special chapter is dedicated to my life partner, Mr. Mohammad Haroon Khan. I cannot begin to express my unfailing gratitude and love to him. Who has supported me throughout this process and has constantly encouraged me when

the tasks seemed arduous and insurmountable. I want to thank him for always standing by me everywhere. It is hard to imagine having made it to this point without him. last but not the least, I wish to thank my angel daughters Ms. Haya Haroon, Ms. Safia Haroon and Ms. Hania Haroon, a source of unending joy and love, who suffered a lot and were ignored during this period.

This thesis is only a beginning of my journey.

At last, I wish to thank many other people whose names are not mentioned here but this does not mean that I have forgotten their help and support.

Abstract

Breast cancer is a multifactorial and complex disorder. It is posing serious public health concerns and its incidence rate is on the rise in Pakistan. It is therefore of prime importance to identify genetic and/or non-genetic factors contributing towards the development and progression of breast cancer. The present investigation is a case-control study including 1000 cases and 1000 age matched controls of the same ethnic background. Individuals were recruited on the basis of a predefined inclusion and exclusion criteria. All participants were in-person directly interviewed after signing an informed consent document. Peripheral blood samples were collected from all the participants along with personal identifiers, demographic characteristics and family history of cancer and other diseases. Vital status/survival status of the patients was determined for up to a maximum of 47 months to record the censored data. We analyzed our sequenced variants and clinico-pathologic features for their possible association with the disease risk by using unconditional logistic regression. Association of the variables was measured with ORs and corresponding 95% confidence intervals. Overall survival of the patients was assessed using Kaplan-Meier curve. Cox proportional hazard model was used to calculate risk ratios and to adjust for potential confounders.

A total of thirteen variants were reported in BRCA1 and BRCA2 genes respectively including three novel variants (Exon3 -37insC, Exon3 -215T<C and Exon14 102-103insTC) in BRCA1 and five novels (exon8 +87insA, exon20 +318T<A, exon19 -351-353delTCT, exon16 -17G<T and exon27 T129A) in BRCA2. Five out of thirteen variants were the in silico identified, HapMap confirmed, pathogenic and previously reported in other populations. Their contribution towards disease risk was tested in our sampled population and it was observed that rs28897686 polymorphism of BRCA1 and rs28897743 of BRCA2 were observed positively associated, while rs28897696 and rs1060915 polymorphisms of BRCA1 and rs4987049 SNP of BRCA2 were found not associated with the disease risk. Five of the eight novel variants, two in BRCA1 (-37insC exon 3 and 102-103insTC exon 14) and three in BRCA2 (+87insA exon 8, -351-353delTCT exon 19 and T129A exon 27)

were observed only in the breast cancer cases and found completely absent in the controls while the rest of 3/8 of the novel variants (BRCA1 -215T<C exon 3, BRCA2 +318T<A exon 20 and BRCA2 -17G<T exon 16) were found highly significantly associated with breast cancer risk. Pairwise Linkage Disequilibrium analysis showed that the strong LD (D′=0.52) exists in between rs28897696 and -215T<C exon 3 variant of BRCA1 and LD (D′=0.43) in between rs28897743 and -17G<T exon 16 of BRCA2.

We also examined the cross-sectional associations of life style, reproductive and socio-demographic risk factors with breast cancer density in Pakistani women. Mean age of cases and controls at recruitment was 50.58±10.68 and 54.78±14.52 years while mean BMI for cases and controls was 26.07±4.04 and 25.05±4.25, respectively. Among the patients 60.70% were married, 46.50% were nulliparous, 16.90% had ≥4 children, 39.90% women breast fed their children, 88.90% were non-smokers and 67.90% were physically active. Post-menopausal women diagnosed with breast cancer accounted for 52.30%. In the current data set, 31.70% patients had at least a blood relative diagnosed with some type of cancer, 22.80% patients were diagnosed with other types of medical complications including high blood pressure, diabetes etc. Significant association between age and breast cancer was observed. Overweight (BMI≥25) and obese (BMI≥30) females have approximately 1.5 times more risk of having breast cancer (Overweight; OR = 1.52, 95% CI: 1.28-1.81 and Obese; OR = 1.41, 95% CI: 1.14-1.74). It was also observed that unmarried women were at more than two fold higher risk. Similarly use of oral contraceptives and smoking were also significantly associated with increasing risk of breast cancer. Individuals who were physically inactive were recorded to be 1.27 times more likely to develop breast cancer. We have found approximately 1.34 fold increase in the disease risk among the postmenopausal patients (OR = 1.34, 95% CI: 1.14-1.58). Breast cancer patients were observed having an overall median survival time of 33 months (95% CI: 28-34).

In this present study we attempted to define the genetic and non-genetic basis responsible for breast cancer incidence among Pakistani population. It can be concluded that there is a significant contribution of BRCA1 and BRCA2 genetic

alterations in breast cancer pathogenesis. It is hoped that our findings will be of great importance to establish adequate evidence-based awareness and preventative measures against breast cancer in Pakistani women.

Keywords: Breast cancer, Pakistani population, BRCA1, BRCA2, Computational analysis, Statistical analysis.

Contents

List of Figures

List of Tables

Abbreviations

APS	Ammonium persulphate
AR	Androgen Receptor
ARE	Androgen Response Elements
BCa	Breast Cancer
BMI	Body Mass Index
CHEK2	Checkpoint kinase 2
CUST	Capital University of Science and Technology
DHQ	District Head Quarter Hospital
EDTA	Ethylene Diamine tetra Acetic acid
EGFR	Epidermal Growth Factor Receptor
ER	Estrogen Receptor
HGMD	Human Gene Mutation Database
HGVD	Human Genetic Variation Database
IBGE	Institute of Biomedical & Genetic Engineering
IRNUM	Institute of Radiotherapy and Nuclear Medicine
95%CI	95% Confidence Interval
OR	Odd Ratio
NORI	Nuclear Medicine Oncology & Radiotherapy Institute
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
PROVEAN	Protein Variation Effect Analyzer
RFLP	Restriction Fragment Length Polymorphism
RR	Risk Ratio
HR	Hazard Ratio

SIFT	Sorting Intolerant from Tolerant
SNP	Single Nucleotide Polymorphism
STE	Sodium Chloride-Tris-EDTA
TBE	Tris/Borate/EDTA
VC	Vice Chancellor
Ensembl	A genome browser
ExAC	The Exome Aggregation Consortium
PolyPhen-2	Polymorphism Phenotyping v2, is a tool which predicts possible impact of an amino acid substitution on the structure and function of a protein
SIR	Standardized incidence ratios
SEER	Surveillance, Epidemiology, and End Results Program
BRCA1	Breast Cancer 1 gene
BRCA2	Breast Cancer 2 gene
PTT	Protein truncation test
HBOC	Hereditary Breast and Ovarian Cancer
CBE	Clinical Breast Exam
MRI	Magnetic resonance imaging
UTR	Untranslated Region
CPS-II	American Cancer Society's Cancer Prevention Study II
FFTP	First Full-term Pregnancy

Chapter 1

Introduction

Cancer is a main source of mortality all around the world [1], [2] and is thus one among the most significant global health problems [3], [4]. Breast cancer is the most widely recognized cancer in females and is becoming the number one killer of women worldwide. It therefore became an increasingly significant question of research all over the world [5]–[7].

Breast cancer (BCa) ranks second among all cancers in women in the developing countries, but its incidence rate is continuously rising and may potentially become the top one in the near future [8]. It is multifaceted issue created by the interaction of hereditary and non-hereditary variables [9]. The probability of developing breast cancer is promoted by multiple factors [10] working together towards the variability in incidence of the disease onset [11].

Breast cancer is caused by a number of risk factors including age, family and personal history of any other cancer type, early menarche, late delivery, abortion either induced or spontaneous, hormonal replacement treatment, late menopause, physical inactivity [12], [13] and most importantly the genetic factors [14]. Hereditary factors contribute for up to ten percent of overall breast cancer cases [15]. BRCA1/2 genes are the most important susceptibility genes for causing breast cancer [16]. Mutations in these two genes alone are so effective to increase the disease risk for 90% of hereditary breast cancer cases [17].

The relationship of BRCA1/2 genes with breast cancer was identified in 1990 and 1994 respectively [18], [19]. Individuals with either BRCA1 or BRCA2 gene mutations may have an expected lifetime risk of about 30-80% [20]–[22], while their germline mutations confer the highest risk and penetrance for breast cancer [23]. Both these genes play their role in multiple processes including cell cycle regulation, DNA double-strand breaks repair, ubiquitination and transcription [24], [25].

Human genome genetic alterations data are rapidly growing due to the introduction of modern technology [26], however, knowledge about their possible disease association and their molecular mechanism is still constrained due to the tedious and relentless nature of exploratory studies [27]. Computational sciences including Bioinformatics can viably produce more significant facts to scale down supplementary exploratory studies [28]. These emerging sciences can likewise select the most enticing cases from the sea of aggregating information. Identification of population specific genetic changes in different racial groups is an imperative stride for the improvement of genetic counseling which makes it promising to utilize a targeted approach for molecular testing. Less expensive procedures can encourage genetic counseling in families with poor financial backgrounds.

The differential inheritance pattern, age of disease onset and phenotypic expression of these two genes point to the possibility that breast cancer risk may also be subjected to modifying genes and other non-genetic factors [29]. Early and late events in life likewise has an impact on breast cancer risk [30], yet it is still exceptionally hard to illuminate why a few females develop breast cancer and others don't [31] which entangles the prevention strategies. Though all women are differentially susceptible to breast cancer, but various hazard elements can increase the likelihood of having the disease. Majority of these factors are irreversible, yet some of them can be modified to decrease the risk [32]. Some groups are at higher risk including those having family history of breast cancer, women with first delivery at older ages [33], those who are using exogenous hormones for longer durations [34] and those facing obesity [35].

A lot of work has been done related to BRCA1/2 protein functions but the understanding is yet incomplete. BRCA1 interacts with various proteins working in DNA repair as signal transducers, damage sensors or repair effectors. It is thus subsequently expected to be persuasive in repair mechanism and genome integrity [36]. Similarly, BRCA2 is dynamically involved in homologous recombination and double strand DNA break repair. In contrast to BRCA1, BRCA2 interacts with only few proteins [37], including RAD51 [38] and PALB2 mediating DNA repair and are also essential for DNA stability and nuclear localization respectively [39]. Both these genes are activated through damage and play their role in various protein complexes for tumor suppression [36].

A large portion of the exploration on BRCA1 and BRCA2 has been conducted on Caucasians, however, higher allelic frequencies of them has been reported in Asians [40] and specifically in the Indo-Pakistani sub-groups [41]. Despite of the fact that Asia resides approximately 60% of the world population, knowledge about hereditary diseases, genetic predisposition and testing is still vague among this population [42].

There exists obvious differences in the occurrence rates and mortality of breast cancer in various regions which suggests that known variables may shift in different parts of the world [43]. The frequency rate of breast cancer has significantly expanded in Pakistani females due to adaptation of westernized life style, genetic profiles, differential risk factor profiles and varying environmental exposures [44]. More than 50% of new cases were diagnosed in women younger than 50 years in Pakistan [45], Singapore [46], Iran [47], Malaysia [48] and Palestine [49]. Pakistani population has comparatively higher rates of breast cancer [50] having a frequency of more than two folds higher as compared to the neighboring countries [51], [52]. According to the Karachi Cancer Registry, the annual age-standardized rate of breast cancer in Pakistan is 69.1 years, which is comparable to European and North American rates [53]. In Asian countries the incidence age for breast cancer is 40-50 years as compared to 60-70 years in the Western countries [54].

1.1 Statement of Problem

Higher allelic frequencies of BRCA1/2 has been reported in Asian population in general [40], [55] and in Indo-Pakistani sub-groups in specific [41]. Knowledge about genetic diseases and its predisposition is still vague among the Asian population, though approximately 60% of the world population resides in Asia [42]. Pakistani population has comparatively higher rates of breast cancer [50] even higher than the neighboring countries [51], [52].

It is unfortunate that majority of literature related to breast cancer has been published based on Western countries and their populations. There exists clear cut variation in the incidence rates and mortality of breast cancer in different geographical regions which suggests that etiological factors differ in their biologic expression and thus have impact on the disease onset [56]. Most of the available figures from developing countries including Pakistan are centered on data from small units due to lack of population based cancer registries [57]. The incidence rate of breast cancer has dramatically increased in Pakistan women which are predicted mainly due to adaptation of westernized life style. It is therefore important to explore Pakistani population for BRCA1/2 mutations/polymorphisms and associated risk factors which might contribute to current knowledge of this vital topic.

1.2 Purpose of Research

Breast cancer is globally one of the leading causes of morbidity and mortality. Genetic aberrations in the BRCA1 and BRCA2 genes are accountable for an augmented risk of breast cancer, however, an inadequate information is available about the relationship of BRCA1/2 genes alterations with breast cancer in Pakistani population. It is therefore needed to inspect the effect of BRCA1/2 genetic changes in breast cancer cases with an emphasis on the anticipated impact of these aberrations on protein structure and function. It is also important to be noted that

because of the differential variation in disease incidence among different ethnicities, etiological factors may also vary and have a profound effect on the disease onset. Exploring the potential breast cancer risk factors in Pakistani women is therefore important, which might contribute to current knowledge of this vital topic. The aims of current study are:

1. To screen Pakistani women for BRCA1 and BRCA2 genes point mutations and polymorphisms in breast cancer
2. To establish the potential pathogenicity of BRCA1/2 genetic alterations in breast cancer in the target population
3. To study the wide ranging functional impacts of genetic alterations of BRCA1/2 genes through sequence, structure and function annotations
4. To study the association of different risk factors with incidence of breast cancer and their impact on survival length of breast cancer patients

This study is hoped to help better understand the potential risk factors and their pattern for breast cancer in Pakistani population which will aid to offer solutions against these factors. This will also aid a real plus to improve the diagnosis and design novel therapeutics against breast cancer in near future.

Chapter 2

Literature Review

This chapter describes a systemic and comprehensive review of the literature. The review of the literature is focused on the prevalence and epidemiology of breast cancer in different populations along with its associated risk factors. The main factors associated with breast cancer risk are the genetic alterations of BRCA1 and BRCA2 genes, demographic, social, clinical, lifestyle and reproductive problems; this was followed by a summary of the findings.

2.1 BRCA1 and BRCA2 Mutations

The Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) assembled data on 18435 families with BRCA1 and 11351 families with BRCA2 mutations from six continents around the globe. They characterised deleterious unique 1650 BRCA1 and 1731 unique BRCA2 mutations from the CIMBA database. Substantial differences were observed in type and frequency of mutations both in ethnicity and geographical distribution. The Consortium also identified a number of mutations that are relatively in high frequency in specific racial groups which can be helpful in designing population based effective strategies against breast cancer [58].

The epidemiology and mutation types in BRCA1 and BRCA2 genes varies depending on population and ethnicity. Mutations in these genes were identified and annotated in the Puerto Rican population including 46 women. All the participants were subjected to genetic testing and Sanger sequencing was performed for the negative individuals on the Hispanel. There were 38 negative genetic testing for BRCA1/2, 4 pathogenic alterations and 4 of uncertain significance. One pathogenic variant (deletion) was observed at exon15-16 in BRCA1 and 3 in BRCA2 [59].

The frequency of germline mutations in BRCA1 and BRCA2 genes were investigated through NGS in 95 unselected Japanese women with ovarian, fallopian tube or peritoneal cancer regardless of their family histories. 12/95 patients had deleterious mutations including 5 in the BRCA1and 7 in the BRCA2. Thirty six cases were having family history and 6 were found having mutations in both BRCA1 and BRCA2. It was also reported that 6/59 cases without a family history also had germline mutations in BRCA1/2. Women with the reported mutation were diagnosed at advanced stages and had poor prognostic histological subtypes. Regardless of a family history, about 13% of the ovarian cancer cases were found associated with an inherited risk, which suggests that BRCA1 and BRCA2 genetic testing should be performed for all ovarian cancers patients [60].

BRCA1/2 mutations are though associated with an enhanced breast cancer risk in males, but only a small amount of data is available on the pathology of male breast cancers (MBCs) carrying BRCA1/2 mutations. A study was conducted to figure out if BRCA1/2 MBCs have a particular obsessive components and whether these contrast from those of BRCA1/2 FBCs. The comparative pathologic features of 419 MBC and 9675 BRCA1/2 FBCs BRCA1/2 mutation carriers were characterized using logistic regression with population based information from the Surveillance, Epidemiology, and End Results (SEER) database. An inverse association of Grade and age was observed among BRCA2 MBCs at diagnosis. Moreover, BRCA2 MBCs were reported to have comparatively higher stage and were probably oestrogen receptor-positive (OR = 10.59, 95%CI = 5.15-21.80) and progesterone receptor-positive (OR = 5.04, 95%CI = 3.17-8.04). Similar pattern

of associations were found among the BRCA1 MBCs and FBCs with the exception of grade. It was also observed that BRCA2 MBCs has comparatively higher grade than MBCs from the SEER database. These results displayed that carriers of BRCA1/2 male breast cancer cases demonstrate distinct pathologic features as compared to BRCA1/2 FBCs [61].

An Irish cohort with breast cancer was investigated for the patterns of BRCA1/2 mutations. Fifty three BCa cases were identified from 1968-2010 among 60 Irish Hereditary Breast Ovarian Cancer (HBOC) families. BRCA1/2 mutations were diagnosed in 50/53 females. Fourteen ladies developed a secondary breast cancer however there was no change in hormone receptor status was observed from primary to secondary cancer. Ten out of fourteen women were included in standard screening, about half of them were diagnosed through mammography, 30% by CBE and 20% by MRI [62].

Breast cancer is very common in Moroccan population, however, the role of BRCA1/2 genes has been to a great extent unexplored. A principal BRCA1 founder mutation c.5309G>T (G1770V) was identified in high-risk five independent Moroccan families through Hereditary testing. Haplotype development was executed to affirm the speculation by utilizing seven BRCA1 microsatellite markers. The examination uncovered a typical haplotype for the studied families, confirming G1770V as the principal founder BRCA1 mutation in the target population [63].

The recurrence of injurious germline mutations was surveyed through NGS by utilizing a panel of 25 genes in a cross-sectional study comprising of 2158 BCa patients. Two cohorts were including in the study, cohort-1 with 1781 patients and cohort-2 of 377 cases. Cohort-1 comprised of patients referred for commercial BRCA1/2 testing while members of cohort-2 were having detailed family/personal history and were previously diagnosed negative for BRCA1/2 alterations. Mutations were most frequently diagnosed in ATM, BRCA1, BRCA2, CHEK2 and PALB2 genes. BRCA1/2 mutations were observed among 9.3% cases in cohort-1 and 2.9% in cohort-2. It was also observed that Ashkenazi Jews has comparatively

lower frequency of mutations as compared to non-Ashkenazi individuals in genes other than BRCA1/2 [64].

An observational study consisted of samples from 33 countries on 6 continents was conducted to explore the role of BRCA1/2 mutations in cancer risk. The study comprised of 19581 carriers of BRCA1 and 11900 of BRCA2 from 55 centers. Hazard ratio was predicted for breast and ovarian cancers on the basis of mutation type, nucleotide position and function. RHR value >1 means elevated breast cancer risk while RHR<1 specifies elevated ovarian cancer risk. 9052/19581 (46%) BRCA1 mutation carriers were diagnosed with BC, 5% with breast and ovarian cancer, 12% with ovarian cancer and 37% were free of any type of cancer. In contrast, 52% of BRCA2 mutation carriers were detected with Breast, 6% with ovarian and 2% with breast and ovarian cancers while 40% were without any malignancy [65].

There are a number of available guidelines which recommends that triple-negative BC females patients must be screened for BRCA1 mutations because of large variations in mutation rate among different populations. A hospital based study was conducted in Mexico City with triple-negative BCa patients to assess the frequencies of BRCA1/2 mutations among 190 women diagnosed at the age $\leq$50 years irrespective of their family history. All the samples were screened for one hundred and fifteen BRCA1/2 recurrent mutations previously been reported in Hispanic women. BRCA mutations were identified in 44/190 cases including fourty three in BRCA1 and only one in BRCA2. A BRCA1 founder mutation ex9-12del of Mexican origin was accounted for 41% of all the detected mutations. On the basis of observed results, it was concluded that BRCA oriented genetic testing carried out for triple-negative cases in the Mexican population [66].

Thirty different types of cancer were identified in a study comprised of 1072 individuals carrying a deleterious BRCA mutation. It was observed that carriers of BRCA1 mutations has a significant increase for breast and ovarian cancer but not for other cancers while carriers of BRCA2 mutations had a significantly higher number of observed vs expected cases in both males and females for pancreatic

(SIR = 21.7, 95%CI = 13.1-34.0) and prostate cancer (SIR = 4.9, 95%CI = 2.0-10.1) [67].

Targeted capture, massively parallel sequencing test called BROCA included all known breast and ovarian cancer genes was used on ovarian cancer probands with a family or personal history of either breast or ovarian cancer to detect the prevalence of mutations. Twenty two probands out of the one hundred and eighteen confirmed probands included in the study were observed with deleterious mutations in BRCA1/2, 27.6% of cases with mutations in ATM, BRIP1, CHEK2, PALB2, PMS2, RAD51D and TP53 genes. Family history of ovarian cancer was diagnosed only for thirty eight patients including 9.8% of ovarian cancer cases having mutations in CHEK2, FAM175A, MSH6, NBN and PALB2 genes. No family history of BCa was in the sampled population. The identification of familial mutations in these patients might be valuable to ascertain the risk of other cancers and to direct therapy [68].

Two hundred and fourteen Bahamian women with invasive breast carcinoma were recruited in a study to ascertain the full range of BRCA1/2 founder mutations. A previously identified founder mutation was found in 49/200 women, one novel founder mutation of BRCA2 in exon-17 (818delA) in 4/200 patients along with five other unique mutations were diagnosed in in BRCA1/2. About 27% of the BC cases were attributed to mutation in BRCA1/2 genes. The observed prevalence rate far exceeds any other country [69].

Genetic testing is globally growing for BRCA mutations which may help to adopt preventive strategies and/or to choose the best chemotherapy. It is essential to have knowledge about the pathologic features of BRCA-associated BCa to develop and deliver personalized treatments. BRCA mutations are reported in different populations all through the globe and it is indispensable that the advantages of hereditary testing and targeted therapies be made available to ladies living outside of western Europe and North America [70].

A cohort of 250 high risk cancer affected Israeli females of different ethnic backgrounds including Ashkenazi were investigated for mismatch mutations. Jewish

women were pre-screened and were affirmed with having no prevalent BRCA1/2 Jewish mutations. Overall, 22/250 women were detected with 10 mutations in BRCA1 and 9 in BRCA2, including three novel highly pathogenic mutations. Three out of the 19 observed mutations were detected in Ashkenazi, 6 in non-Ashkenazi, 2 in non-Jewish Caucasians, 6 in Muslims and 2 in Druze [71].

BRCA1 and BRCA2 mutations were studied in 82 French Canadian families by a group of researchers through DNA sequencing and multiplex ligation-dependent probe amplification assay. Pathogenic alterations were identified in 37 families. Young participants with mutation positive background of hereditary breast cancer (HBC) were significantly more susceptible to BC. The study proposed that BRCA1 and BRCA2 genomic rearrangements are improbable to account for hereditary breast cancer [72].

Germline BRCA1/2 mutations are strong candidates for prediction of BC development. Despite the fact that, the estimation of both mutation penetrance and prevalence is conflicting and notorious, yet are imperative for the comprehension of more precise risk information [73]. An eleven membered Eastern Finnish with multiple BC patients was screened for BRCA1/2 mutations. Five relatives were observed with BRCA2 4088insA mutation, significantly associated with BC and provide comprehensive information related to the role of an individual mutation [74].

Twenty five distinct mutations in 40 individuals including 12 novel mutations were diagnosed in 354 Korean BC patients. A BRCA2 mutation, c.7480C¿T was detected in eight unrelated patients which represents half of the total BRCA2 mutations observed in the target population [75]. Similarly, three novel BRCA2 deletion mutations of familial origin were detected in high risk 312 patients through Multiplex ligation dependent probe amplification method [76].

Ionizing radiations have been reported as one of the key hazard for BCa. It was proposed that, individuals carrying BRCA mutations might be more vulnerable to these radiations. The association of chest x-rays and BCa was investigated in

a group study of 1601 BRCA1/2 women carriers. An increased disease risk was observed in cases older than 40 years [77].

The patterns of pathogenic BRCA1/2 mutations were studied in Czech families with repeated BC history including 96 individuals and 55 patients having no reported BC family history. A total of 35 BRCA1/2 mutations including four in 31 early onset BC females. Seven cases affected with primary BC were identified with a single mutation and 14 patients with medullary breast carcinoma were found with three BRCA1 mutations. Overall, 35/151 cases were detected with BRCA1 mutation and 9 with BRCA2 mutations [78].

Claes and his coworkers explored the effect of splice site variants on splicing mechanism through RT-PCR analysis. A total ten BRCA1/2 variants including eight previously reported and two novel splice site variants were detected in BC families. It was also noticed that BRCA1 4304G_A, IVS3_3A_C and IVS19_2delT and BRCA2 IVS23_2A_G, IVS6_1G_A and IVS24_1G_A lead to altered transcripts and were thus accordingly thought to be responsible for BC development [79].

Pakistani population is at a much higher risk of BC than any other Asian population. A detailed analysis was conducted with 341 BC patients from Karachi and Lahore, Pakistan to scrutinize the impact of genetic factors in cancer development. A 6.7% incidence of BRCA1 and BRCA2 mutation was observed among the subjects. BRCA1 mutations were found 65% of the total identified mutations while most of detected mutations were unique to the target population. Six candidate founder mutations were identified in the study including five in BRCA1 (IVS14-1ArG, 2080insA, 3889delAG, 4184del4 and 4284delAG) and one in BRCA2 (3337CrT). The study presumed that hereditary variables play a critical part in BC development in Pakistan and prevailing BRCA mutations are the principle contributors [80].

Peelen and his coworkers detected 79 BRCA1 mutations in 643 Dutch and 23 Belgian hereditary breast cancer families. They reported 28 novel mutations, including 12 recurrent mutations. 2804delAA mutation was observed nineteen times in the target population which was never been diagnosed previously outside

the Netherlands. A common haplotype spanning approximately 375kb was also identified, representing the occurrence of multiple founder mutations in BRCA1 [81].

One hundred BC families were screened in Helsinki University Central Hospital, Finland to explore hereditary mutations in the coding region and splice boundaries of BRCA2 gene. Five mutations were detected in eight families responsible for premature protein truncation and three alterations were found in multiple families. In haplotype analysis, a common founder mutation was predicted for each recurrent mutation. It was also found that 999del5 recurrent mutation was previously been reported in Icelandic population [82].

2.2 BRCA1 and BRCA2 Polymorphisms

The relationship between BRCA1 polymorphisms and cancer was evaluated in a meta-analysis including thirty five studies and comprised of 28094 cases and 50657 controls. No significant association was observed in between overall cancer risk and rs799917 and rs1799966 polymorphisms in any genetic models. However, the rs16941 polymorphism could significantly increase the overall cancer risk among Caucasian populations only in homozygous and recessive models while rs799917 polymorphism is inversely proportional to the risk of cervical cancer, esophageal cancer, gastric cancer and non-Hodgkin lymphoma among Asian populations. The study also showed that rs1799950 polymorphism decreases the risk of breast cancer in among Caucasian populations [83].

Three tagged missense variants were genotyped on BRCA1 and BRCA2 genes in a total of 603 Chinese patients with pancreatic cancer to explore the correlation between BRCA1/2 polymorphism and disease risk. BRCA1 polymorphism (rs1799966) showed direct association with poor prognosis while the missense variants of BRCA2 (rs766173 and rs144848) showed non significant association with overall disease free survival [84].

A number of SNPs have already been associated with the BC risk at different loci, which accounts for about 10% of the familial cases. A group of researchers screened 548 BRCA1 and 523 BRCA2 mutation carrier females from the Manchester genetic database for mutation and SNP profiling irrespective of their age and disease status. They used Kaplan-Meier curves for Survival analysis and multivariable Cox proportional hazards model for screening patterns of genetic, demographic and clinical variables. Median survival of 46 years was estimated for BRCA1 and 48.9 years for BRCA2 carriers while an average Harrell's concordance index (1-c-index) of 0.221(0.019) for BRCA1 and 0.215(0.018) for BRCA2 carriers was observed. An improved prediction performance was observed for the integrated SNP score coupled with clinical and demographic markers [85].

Mutations and SNPs in both BRCA1 and BRCA2 genes were inspected to ascertain BRCA1 haplotypes in an Indian population. BRCA1 mutations were observed in 52% early-onset BC cases and 57% of their relatives. Ten deleterious alterations were diagnosed in BRCA1 including IVS14+1G>A, IVS17+1G>T, 187delAG, 632insT, Q759X, Q780X, 1052delT, R1203X and 5154delC while five in BRCA2 (W3127X, 4075delGT, 5076delAA and 6079delAGTT). The controls were observed with an enhanced rate of G203A, A3624G and A7470G SNPs in BRCA2 gene which indicates their protective effect [86].

Several BC genetic susceptibility variants were identified in the overall population through GWAS. The pattern of association varied for the observed variants of BRCA1/2 defined by the estrogen receptor status. BRCA1/2 carriers may be among the initial groups for whom an appropriate risk profiling could be developed by using the GWAS identified common BC susceptibility variants [87].

Polymorphisms in the 3'UTR are able to disrupt microRNA binding and may act as predictive risk markers. The BRCA1 3 untranslated region was therefore studied for SNPs in the miRNA binding site in Thai women. The polymorphic alleles showed positive association with familial breast and ovarian cancer [88]. Pelletier and his colleagues sequenced case-control samples (n=221) with known ethnicities and BC subtypes to detect BRCA1 polymorphisms disrupting miRNA

binding. They found three polymorphic variants in 3'UTR of BRCA1 while their haplotype revealed that the cases harbor five rare haplotypes among the controls. BRCA1 3'UTR functional variant (rs8176318) was included in 3/5 rare haplotypes. It was concluded that, these reported SNPs and rare BRCA1 haplotypes may possibly be used as novel predictive markers of BC [89].

The associations of nine SNPs were assessed in a population including 12525 BRCA1 and 7409 BRCA2 carriers. Minor alleles of rs4973768 and rs10941679 were found associated with an increased BC risk in BRCA2 carriers while lack of association was observed for BRCA1 carriers. It was also observed that, rs6504950 has no correlation with BC either in BRCA1 or BRCA2 carriers. Overall, 7/9 SNPs were found associated with BC for BRCA2 and two in BRCA1 carriers (P=0.0049 and 0.03). Based on the mutual genotype distribution, 5% of high risk BRCA2 carriers were predicted with a probability of 80-96% of developing BC by the age of 80. These findings suggested that the observed risk differences are potentially sufficient to effect clinical management of mutation carriers [90].

Multiple BC associated SNPs were identified in the general population through GWAS. Minor alleles of three SNPs, one each in FGFR2, MAP3K1 and TNRC9 were observed to enhance BC risk in BRCA1/2 mutation carriers. Minor allele of rs3817198 was found responsible for increasing the disease risk only in BRCA2 mutation carriers (HR = 1.16, 95%CI = 1.07-1.25). rs13387042 at 2q35 was best fit in dominant model for carriers of both BRCA1 and BRCA2 genes while rs13281615 was found having no relation with BC in either of gene carriers. It was further observed that the 2q35 and LSP1 SNPs interact multiplicatively and enhance the disease risk in BRCA2 mutation carriers [91].

SNPs in FGFR2 (rs2981582), MAP3K1 (rs889312) and TNRC9 (rs3803662) were genotyped in a population of 10358 mutation carriers to examine their relationship with BC risk. Common alleles of these SNPs were found associated with an enhanced BC risk. The rs3803662 SNP of TNRC9 gene was observed associated with the disease risk in both BRCA1 and BRCA2 mutation carriers (HR = 1.13, 95%CI = 1.06-1.20) while minor alleles of rs2981582 and rs889312 were found

inversely correlated with BC risk only in BRCA2 mutation carriers (HR = 1.32, 95%CI = 1.20-1.45 and HR = 1.12, 95%CI = 1.02-1.24) [92].

RAD51 is an important component of double-stranded DNA-repair mechanism, interacting with BRCA1/2 genes during the process. A 5?UTR polymorphism (135G<C) of RAD51 has been proposed as a probable modifier of BC risk in BRCA1/2 mutation carriers. Genotype data of RAD51 (135G<C) SNP was pooled from 19 studies with 8512 female mutation carriers. CC homozygotes were found to increase BC risk (HR = 1.92, 95%CI: 1.25-2.94). When the carriers of both BRCA1 and BRCA2 mutation were independently investigated, only BRCA2 mutation carriers were detected with significantly increased risk for both heterozygotes and homozygotes. It was also explored that the splicing variant alters splicing of RAD51 and thus affect its expression in BRCA2 mutation carriers [93].

2.3 Breast Cancer Risk Factors

There are a number of factors contributing positively towards the increase in breast cancer risk and thus are able to estimate the disease the risk. A case–control study based on Indian population including 100 cases and 101 controls highlighted the risk factors for breast cancer. According to the study, waist size and waist-hip ratio are the major contributing risk factors for breast carcinoma. It was also recommended that adequate weight control and exercise can effectively help to reduce the disease risk [94].

According to a prospective observational study including 74177 females from the Nurses Health Study, an overall weight at the age of 18 years was found inversely correlated with both premenopausal and postmenopausal breast cancer. It was further revealed that, long term gain of weight from the age of 18 during pre-menopause and/or postmenopause were positively associated with postmenopausal cancer risk, while weight gain in premenopause has non significant association with the risk of premenopausal breast carcinoma [95].

Nelson and colleagues investigated low physical activity, post-diagnosis BMI and concurrent diseases to predict BC specific and all-cause mortality. Data was synchronized in After Breast Cancer Pooling Project (n = 9513) from three US female BC survivor cohorts. Positive association of very low physical activity with BC mortality was revealed in an individual lifestyle model and also in the integrated model of all the three lifestyle variables (HR = 1.22, 95%CI = 1.05-1.42). The researchers also observed significant correlation of the three lifestyle variables with all-cause mortality. The strength and significance of associations for comorbidities and very low physical activity remained unaffected in the combined model, while the association with obesity was totally tempered [96].

A population based prospective study was conducted in Wisconsin, New Hampshire and Massachusetts to evaluate the possible relation of smoking status after and before BC diagnosis and mortality. The participants included 20691 women with age ranges from 20-79 years. Out of the total, 6778 women died during a median of 12 years, including 2894 mortalities as a result of BC. Active smokers before disease diagnosis were more probable to die of BC (HR = 1.25; 95%CI = 1.13-1.37), respiratory cancer, cardiovascular and other respiratory diseases, similarly, ladies who kept smoking after disease diagnosis were having higher death probability due to BC (HR = 1.72; 95%CI = 1.13-2.60) [97].

Data from 73388 women in the American Cancer Society's Cancer Prevention Study II (CPS-II) Nutrition Cohort including 3721 invasive BC cases were examined to study the controversial relationship between active smoking and BC. The incidence rate was higher in current (HR = 1.24, 95%CI = 1.07-1.42) and former smokers (HR = 1.13, 95%CI = 1.06-1.21). Ladies who started smoking after or before menarche had a higher hazard. No association was observed for other smoking parameters. Alcohol consumption has no effect on the relationship of smoking status [98].

A Random-effects meta-analyses was carried out as a part of World Cancer Research Fund Continuous Update Project to investigate the pattern and magnitude of relation between obesity and post-breast cancer survival. The study comprised

of 82 studies, including 213075 BC survivors and 41477 deaths. A relative risk of total mortality were RR = 1.10, 95%CI: 0.92-1.31 for underweight, RR = 1.07, 95%CI: 1.02-1.12 for overweight and RR = 1.41, 95%CI: 1.29-1.53 for obese women before diagnosis. Obesity was found associated with higher risk of mortality and poor overall and BC survival in both pre and post-menopausal cases [99].

The correlation between smoking at diagnosis and BC specific and overall survival was examined among 5892 female participants with invasive BC. Current smokers were comparatively had a slightly higher but non-significant BC specific mortality (HR = 1.15, 95%CI: 0.97-1.37). The disease specific mortality in current smokers was observed positively correlated with cumulative exposure to intensity and duration of smoking. Similarly, disease specific mortality found elevated up to 32-56% among heavy smokers. The study also revealed a significant increase (33%) in BC mortality in active smokers at diagnosis compared to never smokers. It is thus evident that both overall and BC specific survival is inversely proportional to smoking at diagnosis [100].

A study was conducted to explore the relationship amongst passive and active smoking, invasive BC risk and possible effect of known risk factors in a population of 322988 women. Present, previous and currently exposed passive smokers vs never smokers were detected with an increased BC risk. The results showed that pack-years increases the disease risk from menarche to first full-term pregnancy (FFTP) (HR = 1.73, 95%CI = 1.29-2.32) for every increase of 20 pack-years. The results suggested that smoking is related to increased BC risk and smoking between menarche and FFTP is particularly damaging [101].

Cross-sectional relationship between weight changes in pre one to post two years diagnosis and functional limitations were explored in a study using logistic regression. Females having BMI greater than 30 kg/m2 had comparatively higher physical limitations. Comparatively more physical limitations were observed in participants with a gain of $\geq$10% of their pre-diagnosis weight (OR = 1.79, 95%CI = 1.23-2.61), a moderate/severe limitation (OR = 2.30, 95%CI = 1.75-3.02) and a lower body limitation (OR = 2.05, 95%CI = 1.53-2.76). Extensive weight loss

in ordinary weight ladies without comorbidity was found related with elevated functional limitations, whereas among obese and/or overweight females, whereas, the same was correlated with a lower risk. Moderate weight loss was found related with an enhanced risk of moderate/severe physical limitation in overweight/obese ladies with comorbidity, whereas, extensive weight gain was though found related with an increased risk but the correlation may depend on comorbidity status and initial BMI [102].

Andersen and his coworkers examined whether mammographic density (MD) affects the associations among birth weight, childhood BMI and height with BC risk. Ladies (n = 13572) with age ranges 50-69 years were followed for BC until 2010 in the Copenhagen mammography screening program. Associations among the target parameters were investigated through logistic and Cox regression models. It was reported that, 8194 ladies had mixed/dense breasts and 716 had developed BC. Childhood BMI was found inversely proportional to mixed/dense breasts at all ages (OR = 0.69, 95%CI: 0.66 to 0.72) at age 7 to 0.56(0.53-0.58) at age 13. No statistical relation was observed between birth weight and mammographic density, height and MD or BC risk and birth weight whereas, height was reported positively linked to breast cancer risk [103].

Obesity has been reported associated with poorer prognosis in early BC, but this relationship depend mainly on the strongly on the positivity of estrogen receptor (ER) and ovarian activity. A study comprised of 80000 patients, only mild association of BMI with BC mortality was observed in 20000 women with ER-poor disease and no association after adjustment for nodal status and tumor diameter. In contrast, BMI was significantly associated with BC mortality in pre/peri and in post-menopausal women (In 60000 women) having ER+ disease, but the association remained significant only in 20000 pre/peri-menopausal women with ER+ disease after adjustment for tumour characteristics. Similarly, when ER+ disease was subdivided by age instead of menopausal status, obesity was noticed significantly relevant only to age about 55 years [104].

Ritte and his colleagues defined the associations among adult height, leg length, sitting height and age at menarche within the European Prospective Investigation into Cancer and Nutrition cohort with breast cancer risk in 990 estrogen receptor and progesterone receptor negative (ER-PR-) and 3,524 ER+PR+ cases through Cox proportional hazards model. It was reported that the risk of both ER-PR- and ER+PR+ breast tumors were directly correlated with leg length, standing height and sitting height but was inversely related to increasing age at menarche. Sitting height had a stronger association (HR = 1.14, 95%CI = 1.08-1.20) with disease risk than leg length (HR = 1.05, 95%CI: 1.00-1.11) in ER+PR+ cases. Tall ladies with an early menarche had up to double risk of developing ER+PR+ tumors but no association was observed for ER-PR- disease. A possible hormonal link may be responsible for the stronger associations between height and ER+PR+ tumors among older ladies that could be specific for post-menopausal cases [105].

Data from two populace based registries was joined to yield information of 16970 parous ladies with intrusive BC. Breast cancer survival was assessed in connection to age at diagnosis, parity, age at first birth and time since last birth through Cox regression. BC survival get decreased with younger age at diagnosis in ladies diagnosed before the age of 50, whereas the survival reduced with older ages at diagnosis. Survival was more awful in ladies diagnosed before the age of 50 years and with at least four births contrasted with ladies having up to a couple of births (HR = 1.3, 95%CI: 1.1-1.6). A brief span since last birth was connected with a decreased survival, but adjustment for grade and stage mitigated the association [106].

A study was intended to look at the relationship between BMI at diagnosis, histopathologic components of BC and frequency of various subtypes by utilizing HER2/neu expression and hormone receptors in a population including 592 premenopausal and 1556 postmenopausal women. Higher BMI in both pre and postmenopausal ladies was discovered associated with bigger tumors. Stout premenopausal patients exhibited poorer histopathologic characteristics when contrasted with under and ordinary weight patients. Postmenopausal ladies with BMI>25 typically developed ER/PgR positive tumors, while no affiliation was

seen in premenopausal cases. Moreover, no association was seen amongst BMI and BC subtypes both in pre and postmenopausal ladies [107].

The Association of reproductive risk factors with BC risk was categorically evaluated in age groups 55-69 and ≥70 years through Cox regression in 58426 Norwegian women followed from 1961-2008 for the incidence of BC. The associations did not vary significantly in the two age groups. Hazard ratios for late vs early age at menarche; late vs early age at first birth; high vs low parity and late vs early age at menopause, were found (HR = 0.79, 95%CI = 0.62-1.01), (HR = 1.54, 95%CI = 1.13-2.11), (HR = 0.68, 95%CI = 0.54-0.86) and (HR = 1.44, 95%CI = 1.10-1.90) respectively in age group ≥70 years. It was thus suggested that reproductive factors may have life-long impacts on BC risk [108].

The association of Age, Obesity and Incident BC Phenotypes were explored for a total of 1001 females diagnosed with invasive BC at Duke University Medical Center. Median BMI of the target population was 27.7 Kg/m^2 while median age was 55.7 years. Increasing BMI was found associated with older age whereas, BMI was almost equally distributed among the age groups. Overall, the ratio of Her2, Luminal B and triple negative subtypes were approximately the same with respect to BMI, however an increasing pattern was observed for luminal A cancer cases with increasing BMI. A lower proportion of the triple-negative and higher proportion of the luminal A phenotypes with older age group were detected after stratification by age. These findings were limited only to females with higher and no trends were found between BMI and incident phenotype. It was further observed that higher BMI in younger ages were related to triple negative and luminal A phenotypes in older age groups. These results suggested that obesity has a differential impacts on tumor progression based [109].

The association between disease specific mortality and age at diagnosis was ascertain in post-menopausal cases with hormone receptor-positive BC. The study analyzed 9766 cases enrolled in Tamoxifen Exemestane Adjuvant Multinational (TEAM) randomized clinical trial between Jan 2001-Jan 2006. The study included 5349 patients in age group <65 years, 3060 in 65-74 years and 1357 in

the age group ≥75 years with a total mortality of 1043 recorded amid the middle follow-up of around 5.1 years. Disease-specific mortality was found increased with age for the age groups 65-74 years and ≥75 years in multivariate analysis. Similarly, BC relapse and other-cause mortality also increased with age [110].

The relationship between BMI and all-cause disease specific mortality was investigated in a prospective study based on menopausal and hormone receptor status among 653 BC Japanese patients. One hundred and thirty six all-cause and 108 disease specific mortality was recorded during the follow-up period. Increased BMI was correlated with higher risk of all-cause death among premenopausal women after adjustment for clinical and confounding factors. BMI ≥25.8 Kg/m^2 was linked with BC specific mortality, BMI <21.2 kg/m^2 with all-cause mortality and disease specific mortality among ER+or PgR+tumor cases. These outcomes suggested that BMI either below or above the normal range are related with higher rates of mortality, particularly among the pre-menopausal or among hormone receptor positive cases [111].

The strength of association of menarche and menopause with BC risk was assessed in a study comprised of an integrated data from 117 epidemiological studies, including 118,964 cases with invasive BC and 306,091 females without the disease. None of them had used menopausal hormone therapy. An increased BC risk was reported either for every year older at menopause or every year younger at menarche. Premenopausal cases were having higher risk of BC than postmenopausal women. The associations were altered by increasing BMI in post-menopausal cases, but did not affected by a women's age, childbearing history, ethnic origin, smoking, alcohol consumption and use of oral contraceptive. The impact of menopause within same age patients and pattern of age at menopause were comparatively stronger for ER+ patients [112].

Paluch-Shimon and coworkers characterized the association between very young age and adverse characteristics of BC at presentation among Israeli women. The cohort was grouped into "very-young" and "less-young" patients. Sixty-one very

young and 94 less-young patients were perceived and their clinico-pathologic features and survival information were analyzed. Mean age at diagnosis was 29.9 and 40.5 years respectively for very young and less young patients. Exceptionally youthful patients had fundamentally more metastatic BC at introduction. Youthful age was not seen as an independent hazard for decreased survival after controlling for stage and tumor grade. The study reasoned that extremely youthful age among Israeli ladies with BC is related with higher stage at determination, unfavorable pathologic attributes and antagonistic result however is not an independent prognostic variable for survival [113].

A meta-analysis was conducted to rate BC risk factors. Obesity, hormone use, alcohol consumption and nulliparity are the mostly reported BC risk factors, present a relatively modest relative risk for BC. Variables identified with past history of neoplastic illness or atypical hyperplasia and components related with hereditary inclination altogether influence BC risk, with relative hazard ranging from 3-200 among premenopausal BRCA mutation carriers [114].

The International Agency for Research on Cancer reported that around 25% of BC cases are because of increasing BMI and inactive lifestyles. The dominant part of epidemiologic reviews demonstrated that women engaged in 3-4 hours/week of moderate to vigorous exercise have a 30-40% decreased risk for BC. Overweight or stout ladies have a 50-250% higher hazard for postmenopausal BC. Liquor utilization additionally enhance the hazard for both pre and postmenopausal BC. Dietary patterns including utilization of high fats, low fiber, low vegetables/fruits and high simple carbohydrates may likewise increase the ailment hazard. The worldwide patterns of expanding BMI and decreasing physical activity may prompt to increasing BC frequency [115].

It is evident from the empirical research reviewed that BRCA1 and BRCA2 are the most important genes causing breast cancer in different world populations. Smoking, obesity, menopausal status, physical activity etc. are some of the more common factors seen in the research to be associated with breast cancer. It is clear from literature review that only a little has been written about breast cancer risk

factors in Pakistani women and thus there is substantial opportunity for screening Pakistani population for BRCA1/2 mutations, polymorphisms and associated risk factors. This thesis therefore addresses an important topic that is relevant to an international audience.

Chapter 3

Methodology

This chapter outlines the research methods and procedures used for this study. It discuss the research design, ethical statement, methods, setting, sample recruitment, instrumentation and data analysis procedures.

3.1 Participants

This population based case-control study was carried out at Capital University of Science and Technology, Islamabad, Pakistan in collaboration with Institute of Biomedical and Genetic Engineering (IBGE), Islamabad, Pakistan. Approval for this study was granted by the Departmental Scientific Committee of the University. It is a hospital-based case-control study, sampling and lab work was conducted in the year 2012-2015, while patients were followed for up to a maximum of 47 months from the date of their registry in the present study to record the censor data for survival analysis.

The cases were histologically/cytopathology confirmed diagnosed breast cancer patients from the hospital registry and admissions at NORI (Nuclear Medicine Oncology and Radiology Institute) Islamabad, Pakistan, DHQ (District Head Quarter Hospital) Rawalpindi, Pakistan and IRNUM (Institute of Radiotherapy and Nuclear Medicine) Peshawar, Pakistan. The inclusion and exclusion criteria for

breast cancer cases and controls were defined in consultation with the supervisor of the thesis, external collaborator at IBGE and medical experts at the hospitals from where samples were collected. The inclusion criteria for cases was i) they must be confirmed cases of breast cancer ii) they should not have suffered from any major chronic illness before breast cancer diagnosis, iii) they should not have administered with long course of any mineral or vitamin supplements for at least the last two years, iv) should not be suffering from severe malnutrition or hepatic disorders, v) had not used menopausal hormone therapy and vi) having no previous history of breast or any other cancer in the last 5 years.

Controls were collected over a similar time frame and were individually matched with patients for their ages (± 5 years) and financial status. Inclusion criteria for the controls included, i) the attendants of patients who were close relatives of the patients, ii) they must not be suffering from any major chronic illness in the past five years, iii) they should not have taken any mineral supplements or vitamin for longer periods of times or during the last two years, iv) they should not be suffering from severe malnutrition or hepatic disorders, v) had not used menopausal hormone therapy and vi) having no previous history of any type of cancer in their life.

3.2 Blood and Tissue Samples

Peripheral blood (5ml) samples were collected from 1000 breast cancer patients along with 1000 age matched controls in blood collection tubes (Becton, Dickinson and Company) containing an anticoagulant (Acid Citrate Dextrose) to study genetic alterations in the sampled population. Each sample tube was properly labeled according to the lab guidelines and were stored at 4° until DNA extraction.

3.3 DNA Extraction

Genomic DNA was extracted from all the samples by using the standard Organic (phenol-chloroform) method [116], followed by ethanol precipitation. The protocol (annexure-1) was slightly modified according to needs and requirements of the study. Each batch comprised of 20 samples, and it took three days to process each batch for DNA extraction with reliable results and large amount of pure DNA.

Each sample was diluted (50 times) by adding 294μl double distilled water to a 6μl DNA sample. Each sample was quantified at 260nm and 280nm wavelength through UV Spectrophotometer (U-3210, Hitachi, Japan). Optical density (OD) ratio for each sample was calculated as:

$$\text{OD} = \frac{\text{Absorbance at 260nm}}{\text{Absorbance at 280nm}}$$

It is recommended that, the Optical Density (OD) must be in between 1.7-1.9 for a good quality DNA. Concentration of DNA in each sample was calculated as,

$$\text{DNA concentration}(\mu l/ml) = \text{Absorbance at 260nm} \times \text{Dilution factor} \times \text{Correction factor}$$

Working solution of 40ng/μl DNA was then prepared from the stock solution by using the following formula,

$$\text{C1V1} = \text{C2V2}$$

3.3.1 Primer Designing for BRCA1 and BRCA2 Genes

Whole gene sequences of BRCA1 and BRCA2 along with full intronic, 5' and 3' flanking regions were retrieved from Ensembl genome browser (http://www.ensem--bl.org/index.html) as pre-requisite for primer designing and selection of restriction

enzymes etc. Ensembl is a joint venture of the Wellcome Trust Sanger Institute and European Bioinformatics Institute (EBI). It was launched in 1999 in response to the imminent completion of the Human Genome Project, producing genome databases for vertebrates and other eukaryotes and making the data openly accessible on the web.

Exon-specific primers harboring intron/exon boundaries were designed for BRCA1 /2 genes by using Primer3 version 0.4.0 (http://frodo.wi.mit.edu) [117]. Primer3 is an open source, widely used software for designing sequencing primers and hybridization probes. Many of its input parameters cab be customized by researcher to make good primers. The primers used are shown in Annexure-2.

3.4 PCR Analysis

Each case-control DNA sample was amplified with respective primers after conditions were carefully optimized for each primer. An optimized PCR recipe was prepared (Table 3.1) and all the contents of a polymerase chain reaction were kept on ice during processing to avoid degradation.

PCR Master Mix (8.5μl) was added to each PCR tube containing 1.5μl of sample DNA to make a 10μl reaction. Each sample was amplified with the following standard cycling conditions in a thermocycler as follows, 95°C for 5 min; 30 cycles; 95°C for 1-2 min, 54-57°C for 2 min, 72°C for 1 min and finally 72°C for 10 minutes for final extension.

PCR products (10μl) were loaded into gel wells by using separate tips for each DNA sample. TAE (Tris/Acetate/EDTA) and TBE (Tris/Borate/EDTA) were used in electrophoresis as they are the most commonly used buffers for nucleic acids electrophoresis. 120 volts of electric current was applied for a mean time depending upon the size of PCR products. After completion of electrophoresis, the DNA molecules were stained in the gel with ethidium bromide to make them

visible. Ethidium bromide is an intercalating agent and makes the DNA fluorescent under UV light, thus each band containing ∼20ng DNA became distinctly visible.

TABLE 3.1: PCR Recipe used for the amplification of BRCA1 and BRCA2 exons.

Reagents	**Stock Conc.**	**Required Conc.**	**Final Conc /reaction**	**Final Conc. for 27 reactions**
ddH_2O			4.6μl	124.2μl
PCR Buffer	10X	1X	1μl	27μl
$MgCl_2$	25mM	1.5mM	1μl	27μl
dNTPs	0.8μ/μl	0.2mM	1μl	27μl
Taq Polymerase	2μ/μl	0.8μ/μl	0.3μl	8.1μl
Forward Primer	20μM	600nM	0.3μl	8.1μl
Reverse Primer	20μM	600nM	0.3μl	8.1μl

3.5 BRCA1 and BRCA2 SNP Analysis

Single nucleotide polymorphisms (SNPs) represent the most abundant and recurrent type of genetic alterations in the human genome, their patterns probably influence numerous phenotypes. SNP genotyping are consequently anticipated on large scale to identify genes affecting complex diseases/disorders. We used PCR-RFLP method for genotyping of previously reported SNPs in our study, while direct sequencing for the rest of experiments. The required endonucleases for PCR-RFLP experiments were selected by using WATCUT (http://watcut.uwaterloo.ca/watcut /watcut/template.php), an online tool for SNP-RFLP analysis.

Previous association studies of BRCA1/2 variants and BC generally has focused on the alterations in coding regions. However, it is also probable that genetic alterations in the non-coding regions may also be able to influence cancer predisposition [118].

3.5.1 BRCA1 and BRCA2 Coding SNPs

SNPs in the coding region may result in a conformation change, affecting protein functions while promoter polymorphisms may results in change of transcription activity and levels of protein expression [119]. The BRCA1 SNP Leu871Pro, located in the RAD51 interaction domain, was reported to be associated with BC risk [120]. Similarly, the BRCA2 genotype of the N372H polymorphism was found poitively associated with an inrease of overall cancer risk [121]–[123], ∼1.3–1.5 fold increased risk in the Caucasians [124]–[126].

Majority of the studies had been based on polymorphisms identified during BRCA screening for mutations from high-risk families or early-onset breast cancer, and on Caucasian populations. No previous hypothesis driven systematic study of BRCA SNPs relating altered function to risk association has been reported to the best of our knowledge.

It was hypothesized that, low penetrant alleles of polymorphic variants of the BRCA1/2 genes may contribute towards breast cancer risk in Pakistani population by affecting transcriptional and/or functional activities of these genes. Our objective was first to identify potential functional SNPs available from the public databases as well as those reported in literature specific to Pakistani population. These SNPs would first be genotyped in our case-control samples for allele frequency estimation and selection of appropriate SNPs for risk association and haplotype studies.

3.5.2 Identification and Selection of BRCA1/2 SNPs from dbSNP

Functionally important SNPs of the BRCA1 and BRCA2 genes were identified from the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/) through predefined selection criteria. dbSNP database was established by National Center for Biotechnology Information (NCBI) in in collaboration with the National

Human Genome Research Institute (NHGRI) in response to a need for a general catalog of genomic variations which are of prime importance for gene mapping, association studies and phylogenetics. The data within dbSNP is freely and publicaly available in a variety of forms since its establishment in 1998 [127].

According to the pre-set criteria, we included non-synonymous SNPs in the coding region, 100bp of introns on each side of exons, promoter and 100bp downstream of the gene.

3.5.2.1 Identification of BRCA1 SNPs

At the time the study was initiated, in silico search from dbSNP public database identified a total of 5741 SNPs in BRCA1 gene including 1366 missense, 288 nonsense, 63 in 5'UTR, 149 in 3'UTR, 37 in 5' splice site, 38 in 3' splice site and 3800 others. Out of the total 5741 SNPs of BRCA1, 1691 are either missense or nonsense or present in the 5' splice site or 3' splice site or in the promoter region and thus are significantly more important. We restricted our search by selecting only pathogenic or likely pathogenic SNPs (331) from those 1691. It was further confirmed that out of the 331 SNPs only 4 (rs28897686, rs28897696, rs386576387 has merged into rs28897686 and rs386576392 has merged into rs28897696) have been confirmed through the HapMap project and were thus selected for further analysis.

3.5.2.2 Identification of BRCA2 SNPs

A total of 7276 SNPs were identified in BRCA2 gene through in silico search from dbSNP including 2378 missense, 347 nonsense, 42 in 5'UTR, 97 in 3'UTR, 33 in 5' splice site, 32 in 3' splice site and 4347 others. 2771/7276 was comparatively more significant because of their presence in either 5' splice site or 3' splice site or in the promoter region or they were either missense or nonsense in nature. We further restricted our search to only pathogenic or likely pathogenic SNPs, which resulted only 340 SNPs. It was further confirmed that only 2/340 (rs4987049 and

rs28897743) has been confirmed through HapMap project and were thus selected for further analysis.

3.5.2.3 Genotyping of BRCA1 and BRCA2 Selected SNPs

Genotyping of all the selected SNPs (4 in BRCA1 and 2 in BRCA2) was performed on 1000 breast cancer cases and 1000 age matched controls through PCR-RFLP. As the selected SNPs were in the exonic regions so the already designed primers were used to amplify the exonic regions having their respective SNPs.

3.5.3 Statistical Analysis

Allelic and genotypic frequencies were calculated for each genetic variant. Odd ratios (OR) with 95% confidence intervals (95%CI) were used to measure the strength of association between the observed polymorphism and disease risk by using un-conditional logistic regression models in cases-control groups. Regression techniques are versatile and are commonly used in medical research, having the abiliy to predict outcomes, measure associations and has the control for confounding variable effects. Logistic regression is an efficient and powerful technique of multivariate analysis which is most widely used in epidemiology. It is applied to analyze the effect of a group of independent variables on a binary outcome by quantifying each independent variable's unique contribution, allowing the measure of association between the occurrence of a qualitative dependent variable (event) and factors susceptible to influence the event.

Different genetic models including Dominant, Co-dominant, Over-dominant, Recessive and Log additive were used to evaluate the risk and allelic associations while P-value <0.05 was used as a threshold for statistical significance. Kaplan Meier survival analysis was performed to examine the relationship between genotypes and survival using RStudio. In clinical trials or community trials, researchers are usually interested in the time until participants of the study present an event or endpoint including death etc. Kaplan-Meier estimator is one the best option as

a non-parametric statistical method to measure the fraction of subjects living for a certain period of time after being treated.

3.5.4 DNA Sequencing

All the DNA samples from cases and controls were directly sequenced for BRCA1/2 exons with the same forward and reverse primers used in the previous step for PCR amplification at the Institute of Biomedical and Genetic Engineering, Islamabad, Pakistan. PCR products were treated with Exo-I and SAP before sequencing to remove the traces of dNTPs and primers if any. DNA sequencing reactions were carried out through BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The mixtures for cycle sequencing contained 0.5μL of Ready Reaction Premix, 1.75μL of BigDye Sequencing Buffer, 100 ng of purified PCR product, 1.6 pmol of forward or reverse primers and double distilled water adjusted to 10μL in total volume. The reactions were executed under the following conditions: 1 cycle of 96°C for 1 minute, 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes. The reaction products (10μL) were precipitated by adding 1.0μL of 125 mM EDTA (pH 8.0), 30μL of 95% ethanol and 1.0μL of 3M sodium acetate (pH 5.2) and incubated for 15 minutes at room temperature. The mixture was centrifuged at 20000 xg at room temperature for 20 minutes. The DNA pellet was washed with 70% ethanol, centrifuged at 20000 xg for 10 minutes at room temperature, then air-dried and dissolved in 12μL of Hi-dye formamide. The products were then transferred to a 96-well plate and sequences were read in ABI 3700 DNA Analyzer. BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) was used to compare sequence results with reference sequences. The products were then analyzed through SeqScape 2.5 software (Applied Biosystems) and other freely available softwares including FinchTV etc. The sequence data obtained was then analyzed through different Bioinformatics techniques for their in silico characterization and annotations.

3.5.5 Pathogenicity Prediction

Exploration of genetic changes responsible for causing human genetic disorders are assumed to be among the most significant research challenges [128]–[131]. A number of computational approaches are available to predict the potential pathogenicity of diverse types of genetic alterations. Some of these methods are exclusively based on sequence level information and multiple sequence alignments (MSAs) [129], [132], while others on evolutionary information or varying protein structural descriptors. Possible pathogenicity was predicted for all the observed genetic variants in this study by using computational biology approach and exploiting the publicly available servers to examine the likelihood of their structural and functional impacts.

We considered the Breast cancer related mutations and associated SNPs of BRCA1 and BRCA2 genes in Pakistani population to annotate them for their ability and intensity to affect structures and functions of the respective proteins. It was thus hoped that this study will help to enhance our understanding towards the prospective molecular causes and guide to develop new treatment and management strategies against genetic disorders. Systematic computations on the basis of sequences and structures were applied with the help of Bioinformatics techniques to study the effect of genetic variations as they affect the protein structures and other physiochemical properties, thus damaging their interactions.

Biochemical assays, though are used globally to identify deleterious alterations, but are time consuming and laborious. Accurate and precise computational analysis of the possible functional outcomes of genetic variants can significantly decrease the time complexity by ranking them probably to be deleterious or vice-versa. It is thus believed that computational methods, able to categorize and propose descriptions for the possible impact of genetic variants can complement functional tests. In our study we used automated methods i.e. SIFT [133]–[136] to predict the pathogenicity of mutations. These tools are capable of discriminating between driver and passenger mutations.

3.5.6 Mutation Analysis

BRCA1 and BRCA2 genes are autosomal dominant and of high penetrance. Both of them are the most commonly implicated genes in breast cancer. Genetic alterations of both these genes can significantly increase the risk of breast and ovarian cancers. Germline mutations of these genes are estimated to account for 5-10% of all BC and approximately 80-90% of all hereditary BCs [137]. The cumulative risk of BC with a BRCA1 mutation is about 3% by age of 30, 19% by age of 40, 51% by age of 50, 54% by age of 60 and 85% by age of 70 years [138]. Similarly, BC risk in BRCA2 mutation carriers was estimated as 32% by age 50, 67% by age of 70 and 80% by age of 90 years [139].

Hundreds of BRCA1/2 mutations have already been reported in the Breast Cancer Information Core database (http://research.nhgri.nih.gov/bic/). We annotated our reported mutations by exploiting computational biology approaches including sequence, structure and functional annotations, pathogenecity establishment, Linkage disequilibrium analysis etc. With the aid of bioinformatics and computational biology approaches, one can model mutations reasonably and quickly to investigate the effects of genetic changes on protein structure and function well before the experimental characterization of engineered proteins.

3.5.7 Haplotypes and Linkage Disequilibrium Analysis

Haplotype effects were predicted for haplotype-based analyses of both BRCA1 and BRCA2 genes polymorphic variants among all samples irrespective of BRCA1/2 mutation carrier status. Haplotype analyses were conducted through online SNPStats tool [140], which is a simple and ready to use software package of R, specially designed for the analysis of genetic epidemiology and SNP association studies. Expectation maximization (EM) algorithm was used for the prediction of haplotype frequencies while the most common haplotype in the study was used as a reference group. Odd Ratios and 95% CIs were calculated through an unconditional regression under additive model. Haplotypes with frequency of <3% were ignored in

the study. Pairwise linkage disequilibrium (LD) measures (D' value) and linkage disequilibrium plots were generated through SHEsisPlus [141], [142], a web-based multi-purpose platform. The LD coefficient $'D'$ has a range of possible values depending upon the frequencies of alleles and thus is not a convenient measure of LD. D' was therefore used instead of $'D'$, using the theoretical maximum of $'D'$ for its normalization. D' ranges from 0 to 1. $D' = 0$ means no LD, whereas D'=1 means complete LD. Stronger the value of D', stronger will be Linkage disequilibrium.

3.6 Socio-demographic and Reproductive Risk Factors

Cancer is a complex multistep disorder, resulting from a combination of factors [143]. The associated etiological factors of breast cancer including age, environmental, lifestyle factors and reproductive factors are poorly investigated especially in Pakistani population. We, therefore, systematically collected detailed information about all the possible risk factors including demographic, reproductive and lifestyle characteristics of each candidate through a specially designed proforma (annexure-3). An advantage of our study over the previous studies is that it was specifically designed to investigate the association of each factor independently and in combination with other as well.

Literature survey was systematically carried out to have a deeper insight towards the contribution of different risk factors. All participants were in-person asked (direct interview) with similar questions after signing an informed consent document prior to being interviewed. A total of 1000 cases were included in the analysis along with 1000 controls, same individuals used for previous analysis. The data included personal identifiers, demographic characteristics and family history of cancer and other diseases. The vital status/survival status of study participants were determined through the hospital records and making personal telephonic inquiries each month after interview. The participant cases were followed for a maximum of 47 months to record the censored data.

3.6.1 Statistical Analysis

Efforts were made to frequency match the cases and controls by age, ethnicity and geographical location. We analyzed this data to evaluate the clinic-pathologic features of breast cancer patients in the local Pakistani women. The population characteristics of different variables for both the cases and controls and pathologic features for cases were analyzed using R version x64 3.1.0 [144]. Potential association of various clinical/socio-demographic features with breast cancer was measured through ORs and corresponding 95% confidence intervals by using Logistic regression. Overall survival of the patients after diagnosis was assessed using Kaplan-Meier curve.

Kaplan-Meier curves were designed in 1958 by Edward L. Kaplan in collaboration with Paul Meier to deal with incomplete observations. These curves have become a popular method of dealing with differential times-to-event especially if all the subjects do not continue for the whole study [145].

We used Cox proportional hazard model to calculate risk ratios and to adjust for potential confounders. For all of the models, we stratified on age groups, number of kids, menopausal status, physical activity and smoking. BMI were included in all models. The statistical significance of the interaction was evaluated at $P \leq 0.05$ level by using the likelihood ratio test, wald test and logrank test.

Chapter 4

Results and Discussion

This chapter summarizes the study findings of mutations, polymorphisms and factors associated with an enhanced risk of breast cancer in Pakistani women. The results are presented under different headings and sub-headings.

The present study comprised of 1000 cases and 1000 control samples collected from NORI, IRNUM and DHQ Rawalpindi, Pakistan. Nucleotide sequences of BRCA1 (ENST00000357654.7) and BRCA2 (ENST00000380152.7) genes were retrieved from Ensembl (http://www.ensembl.org/index.html) as a pre-requisite exon-specific primers designing and comparative studies. The primers used in this study are shown in Annexure-2.

BRCA1 has 24 exons, covering a length of about 100kb of genomic DNA and encodes a protein of 1863 residues. BRCA2 is larger than BRCA1 gene, consisting of 27 exons and spans about 70kb. It encodes a protein of 3418 residues (Figure 4.1). BRCA1 and BRCA2 proteins have various functions including DNA repair, transcriptional and cell cycle checkpoints regulation.

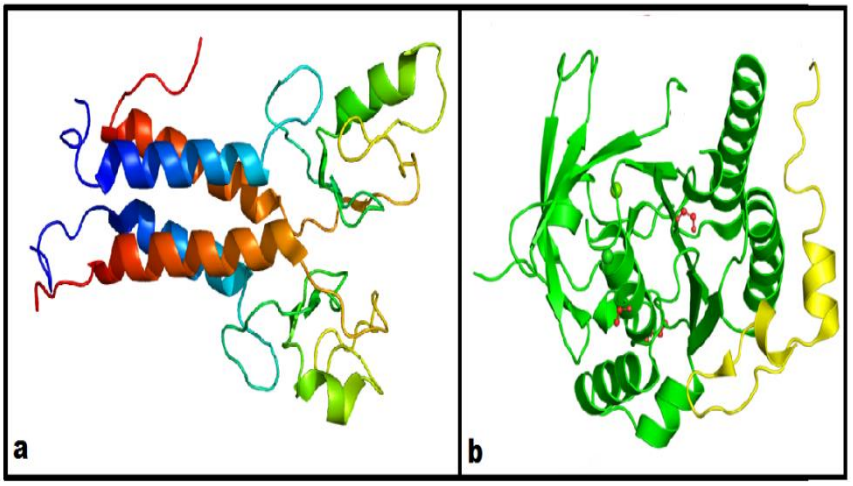

FIGURE 4.1: Crystal structures of BRCA1 and BRCA2
(a) BRCA1/BARD1 RING-domain heterodimer (PDB id=1JM7).
(b) RAD51-BRCA2 BRC repeat complex (PDB id=1N0W).

4.1 BRCA1 and BRCA2 SNPs

One of our objective was first to identify possible functional SNPs available from the public database which were reported pathogenic and also been confirmed through the HapMap project. These SNPs would first be genotyped in case-control study for their allele frequencies and possible association with the disease risk.

We explored the contribution of in silico identified pathogenic and HapMap confirmed previously reported 4 SNPs (rs28897686, rs28897696, rs386576387 and rs386576392) of BRCA1 gene in 1000 breast cancer cases along with an equal number of age matched controls. From the literature and dbSNP, it was observed that the rs386576387 SNP has been merged into rs28897686 and rs386576392 has been merged into rs28897696, so we have actually only two SNPs for genotyping analysis. The logical reason for merging of SNPs in databases is that, if and only when two SNPs map to the same contig/position and have the same variation class are merged into one in a subsequent build by the database curators.

Similarly 2/340 in silico identified pathogenic and HapMap confirmed previously reported SNPs (rs4987049 and rs28897743) of BRCA2 were selected for genotyping

analysis. As the selected SNPs are in the exonic regions, the already designed primers were used to amplify the regions having their respective SNPs.

4.1.1 Genotyping of BRCA1 rs28897696 and rs28897686 SNPs

BRCA1 rs28897696 (A1708E) SNP (exon 16) was genotyped in a case-control study of 1000 breast cancer patients along with an equal number of age matched controls through PCR-RFLP. Amplification was carried out with Forward primer 5'-TCTTTAGCTTCTTAGGACAGCACTT-3' and Reverse primer 5'-CTCAGCAT CAGCAAAAACCTT-3' wile restriction digestion with Ssi1 endonuclease. The PCR product is 250bp length, was digested overnight with the restriction enzyme Ssi1. The restriction enzyme Ssi1 cuts in the presence of major allele C. If C is present in homozygous form it will be digested into two fragments of 124 and 126bp, minor allele A will remain uncut and will produce a single fragment of 250bp, while heterozygous CA will produce three bands (124 + 126 and 250bp). Restriction fragments were then analyzed on 4% agarose gel stained with ethidium bromide and were visualized under UV light in gel documentation (Figure 4.2). Allele and genotype frequencies of the rs28897696 polymorphisms for breast cancer patients and control groups are summarized in Table 4.1. There was no significant association of the genotypes in any of the genetic model with increased breast cancer risk (Table 4.2).

PCR-RFLP analysis were used to genotype the rs28897686 (G/A) BRCA1-E1250K polymorphism in exon-10 in 1000 breast cancer patients along with age matched 1000 controls. This SNP is missense in nature and due to the transition of G>A (GAG<AAG), it results in the substitution of E>K at position 1250. The fragment of exon-10 containing the SNP was amplified with forward primer 5'-AGGCATAGCACCGTTGCT-3' and reverse primer 5'-TCTTCCAATTCACTGCA CTG-3'. The resulting 188bp PCR product was digested with restriction enzyme Hpy188I and the fragments were resolved on a 4% agarose gel containing ethidium bromide. The wild-type homozygous G allele produced two fragments of 167

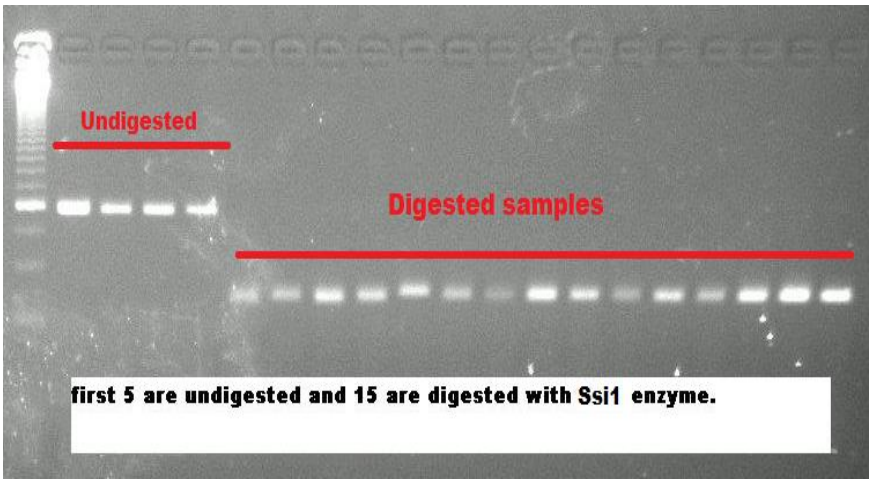

FIGURE 4.2: The 250bp PCR product digested with Ssi1. Homozygous C allele produces two fragments of 124 and 126bp, homozygous A allele remains undigested and produced a single fragment of 250bp, while heterozygous CA produced three bands (124 + 126 and 250bp).

TABLE 4.1: Allele and genotype frequencies of rs28897696 and rs28897686 SNPs of BRCA1 gene in case-control population.

rs28897696 (A1708E) allele frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Allele/Genotype	**Count**	**Proportion**	**Count**	**Proportion**	**Count**	**Proportion**
C	2497	0.62	1360	0.6	41	0.04
A	1503	0.38	640	0.53	548	0.55
A/A	582	0.29	171	0.17	411	0.41
C/A	339	0.17	298	0.3	41	0.04
C/C	1079	0.54	531	0.53	548	0.55
Allele/Genotype	**rs28897686**					
G	2502	0.63	1373	0.69	1129	0.56
T	1498	0.37	627	0.31	871	0.44
G/G	851	0.43	506	0.51	345	0.34
G/T	800	0.4	361	0.36	439	0.44
T/T	349	0.17	133	0.13	216	0.22

and 21bp, the polymorphic homozygous A allele produced a single band of 188bp, whereas the heterozygous GA produced all the three bands of 188, 167 and 21bp. This SNP has been previously described in Spanish [146] and Chilean populations [147].

The genotype frequencies, OR and 95%CI were calculated using R 3.1.1 statistical

computing software [144]. The polymorphic allele showed positive association with breast cancer risk in all the genetic models, highest risk in co-dominant model (OR = 2.38, 95%CI = 1.84-3.08) (Table 4.2).

TABLE 4.2: Association between rs28897696 and rs28897686 SNPs of BRCA1 gene and breast cancer risk.

Model	**rs28897696 (A1708E) (n=2000)**				
	Genotype	**Controls**	**Cases**	**OR (95% CI)**	**P-value**
Co-dominant	C/C	171 (17.1%)	411 (41.1%)	1.00	<0.0001
	C/A	531 (53.1%)	548 (54.8%)	0.43 (0.35-0.53)	
	A/A	298 (29.8%)	41 (4.1%)	0.06 (0.04-0.08)	
Dominant	C/C	171 (17.1%)	411 (41.1%)	1.00	<0.0001
	C/A-A/A	829 (82.9%)	589 (58.9%)	0.30 (0.24-0.36)	
Recessive	C/C-C/A	702 (70.2%)	959 (95.9%)	1.00	<0.0001
	A/A	298 (29.8%)	41 (4.1%)	0.10 (0.07-0.14)	
Over-dominant	C/C-A/A	469 (46.9%)	452 (45.2%)	1.00	0.45
	C/A	531 (53.1%)	548 (54.8%)	1.07 (0.90-1.28)	
Log-additive	—	—	—	0.29 (0.24-0.33)	<0.0001
rs28897686 (n=2000)					
Co-dominant	G/G	506 (50.6%)	345 (34.5%)	1.00	<0.0001
	G/A	361 (36.1%)	439 (43.9%)	1.78 (1.47-2.17)	
	A/A	133 (13.3%)	216 (21.6%)	2.38 (1.84-3.08)	
Dominant	G/G	506 (50.6%)	345 (34.5%)	1.00	<0.0001
	GA-A/A	494 (49.4%)	655 (65.5%)	1.94 (1.62-2.33)	
Recessive	G/G-G/A	867 (86.7%)	784 (78.4%)	1.00	<0.0001
	A/A	133 (13.3%)	216 (21.6%)	1.80 (1.42-2.28)	
Over-dominant	G/G-A/A	639 (63.9%)	561 (56.1%)	1.00	4e-04
	G/A	361 (36.1%)	439 (43.9%)	1.39 (1.16-1.66)	
Log-additive	—	—	—	1.59 (1.40-1.80)	<0.0001

BRCA1 gene was found to lie in a region of extensive linkage disequilibrium, which makes it possible to test a subset of variants in this region representing all the common genetic alterations [148]. Dunning and his colleagues genotyped four missense SNPs in breast and ovarian cancer affected white women and identified three haplotypes, but none of these haplotypes were significantly associated with cancer risk [149]. Cox and his coworkers genotyped 4 tagging SNPs and identified five common haplotypes with a frequency $\geq$5% by using sequencing data. One of the five haplotypes was observed correlated with an enhanced breast cancer risk (OR = 1.18, 95%CI = 1.02-1.37) [150]. However, a later study examining nine tagging SNPs in about 900 Caucasian participant, showed non-significant association of any haplotypes effecting breast cancer risk [151]. Similar results were reported by another group of researchers i.e. lack of association between breast cancer risk and BRCA1 haplotypes [152].

4.1.2 Genotyping of BRCA2 rs4987049 and rs28897743 SNPs

Both the BRCA2 SNPs (rs4987049 and rs28897743) were genotyped in 1000 cases along with an equal number of controls. The rs28897743 (R2336H) SNP is located at the consensus splice donor site in exon13 and has been shown to disrupt normal splicing. It results in two splice variants, one skipping exon13 and the other skipping exons12 and 13. Both the transcripts result in a frame shift and generating a premature stop codon in exon14 [153]. The rs28897743 polymorphism is located in the FANCD2 and FANCG binding domains of BRCA2 [154]. It is therefore possible that it may affect the interaction of BRCA2 with these proteins, which may lead to disrupt the organization of the DNA repair complex.

Using 21 tagging SNPs, Freedman and co-workers conducted a haplotype analysis of BRCA2 with breast cancer risk in a Multi-Ethnic Cohort [151] and found a haplotype tagged by SNP rs206340 significantly associated with an increased risk of breast cancer in homozygous cases (OR (AA/GG)=1.59; 95%CI=1.18-2.16). However, Baynes and his colleagues reported an equivalent tag SNP rs206343 with

a non-significant trend for the minor allele (OR (GG/AA)=0.89; 95%CI=0.73-1.09) [152].

Both the SNPs of BRCA2 gene (rs4987049 and rs28897743) were investigated for their possible association with the disease risk. The rs4987049 SNP was observed not associated with the disease risk in any of the tested models except the Log-additive model (OR=2.90, 95%CI=1.03-8.12), while rs28897743 was in significant association with breast cancer risk in Pakistani population under all the genetic models (Table 4.3 and Table 4.4).

TABLE 4.3: Allele and genotype frequencies of rs4987049 and rs28897743 SNPs of BRCA2 gene in case-control population.

rs4987049, (TAC ⇒ TAG) Y3276X allele frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Allele/Genotype	**Count**	**Proportion**	**Count**	**Proportion**	**Count**	**Proportion**
C	3981	1	1996	1	1985	0.99
G	19	0	4	0	15	0.01
C/C	1984	0.99	996	1	988	0.99
C/G	13	0.01	4	0	9	0.01
G/G	3	0	0	0	3	0
Allele/Genotype	**rs28897743, (CGC ⇒ CAC) R2336H**					
G	3936	0.98	1990	1	1946	0.97
A	64	0.02	10	0	54	0.03
G/G	1949	0.97	992	0.99	957	0.96
G/A	38	0.02	6	0.01	32	0.03
A/A	13	0.01	2	0	13	0.01

TABLE 4.4: Association between rs4987049 and rs28897743 SNPs of BRCA2 gene and breast cancer risk.

rs4987049 association with response STATUS (n=2000, crude analysis)					
Model	**Genotype**	**Controls**	**Cases**	**OR (95% CI)**	**P-value**
Co-dominant	C/C	996 (99.6%)	988 (98.8%)	1.00	0.046
	C/G	4 (0.4%)	9 (0.9%)	2.27 (0.70-7.39)	
	G/G	0 (0%)	3 (0.3%)	NA (0.00-NA)	
Dominant	C/C	996 (99.6%)	988 (98.8%)	1.00	0.04
	C/G-G/G	4 (0.4%)	12 (1.2%)	3.02 (0.97-9.41)	
Recessive	C/C-C/G	1000 (100%)	997 (99.7%)	1.00	0.041
	G/G	0 (0%)	3 (0.3%)	NA (0.00-NA)	
Over-dominant	C/C-G/G	996 (99.6%)	991 (99.1%)	1.00	0.16
	C/G	4 (0.4%)	9 (0.9%)	2.26 (0.69-7.37)	
Log-additive	—	—	—	2.90 (1.03-8.12)	0.021
rs28897743 association with response STATUS (n=2000, crude analysis)					
Co-dominant	G/G	992 (99.2%)	957 (95.7%)	1.00	<0.0001
	A/G	6 (0.6%)	32 (3.2%)	5.53 (2.30-13.28)	
	A/A	2 (0.2%)	11 (1.1%)	5.70 (1.26-25.79)	
Dominant	G/G	992 (99.2%)	957 (95.7%)	1.00	<0.0001
	A/G-A/A	8 (0.8%)	43 (4.3%)	5.57 (2.61-11.91)	
Recessive	G/G-A/G	998 (99.8%)	989 (98.9%)	1.00	0.0086
	A/A	2 (0.2%)	11 (1.1%)	5.55 (1.23-25.10)	
Over-dominant	G/G-A/A	994 (99.4%)	968 (96.8%)	1.00	<0.0001
	A/G	6 (0.6%)	32 (3.2%)	5.48 (2.28-13.16)	
Log-additive	—	—	—	3.75 (1.99-7.05)	<0.0001

4.2 BRCA1 and BRCA2 Sequencing Results

BRCA1 and BRCA2 are autosomal dominant genes with high penetrance and known to play important roles in breast cancer risk. Germline mutations of both these genes represent around 80-90% of all hereditary breast cancers. Their tumor suppressor function in the development of sporadic breast cancer remains unclear, with only a handful of somatic mutations identified in sporadic breast cancer [155], [156].

Sequencing of our samples revealed a total of thirteen genetic variants including eight novel variants along with five previously reported SNPs in BRCA1 and BRCA2 respectively. Each of the reported variant has been individually discussed in detail as follows.

4.2.1 BRCA1 Variants

4.2.1.1 BRCA1 -37insC exon-3 (g.4311816)

This insertion of nucleotide C at -37 position of exon-3 (g.4311816) is a novel variant of BRCA1 gene, confirmed through HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project. This variant was observed only in homozygous condition in twenty two cases and was found completely absent in the control group. These results suggests that the variant has a direct link with the disease development but it still needs further large genotyping studies to clarify its position in the target population (Figure 4.3).

4.2.1.2 BRCA1 T123C at exon-12 (g.43082453)

This transition has been previously reported by other authors (dbSNP id = rs1060915), is synonymous in nature and thus the wild residue Serine at position 1436 remains unchanged. This residue is located in a domain that is important for binding of other molecules. The residue (Ser1436) is part of an interpro domain named

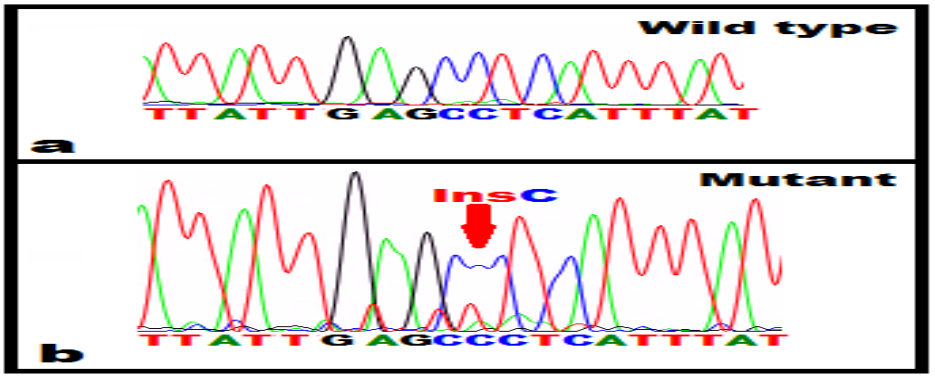

FIGURE 4.3: Novel variant in BRCA1 intron 2 (g.4311816), -37insC exon-3 detected in breast cancer patients in the homozygous state, represented by the red arrow.

BRCA1-Associated (IPR031099), an interpro domain named Breast Cancer Type 1 Susceptibility Protein (BRCA1) (IPR011364).

This domain is annotated with the Gene-Ontology (GO) terms DNA Binding (GO:0003677), Ubiquitin-Protein Transferase Activity (GO:0004842) and Zinc Ion Binding (GO:0008270) to indicate its function. These GO annotations indicate that the domain has its functions in: Nucleic Acid Binding (GO:0003676), Transferase Activity (GO:0016740) and Ion Binding (GO:0043167) (Figure 4.4). No association of the variant was found in any of the genetic models tested (Figure 4.4, Table 4.5).

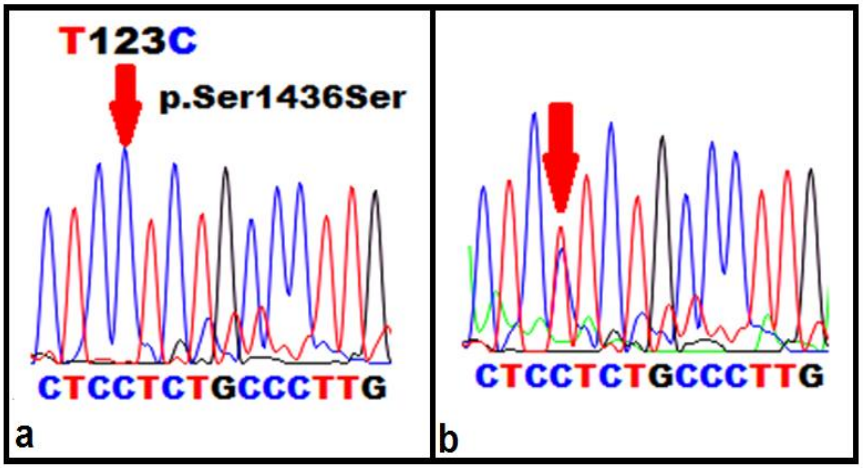

FIGURE 4.4: Identification of Synonymous variant of BRCA1 NC_000017.11(BRCA1_v001):g.43082453T>C (NP_009231.2:p.Ser1436Ser) T123C at exon-12 (dbSNP id: rs1060915), represented by the red arrow (a) Homozygous C (b) Heterozygous TC.

TABLE 4.5: Genotype frequencies and association of BRCA1 gene rs1060915 SNP with disease risk.

rs1060915 allele frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Allele	Count	Proportion	Count	Proportion	Count	Proportion
T	2765	0.69	1317	0.66	1448	0.72
C	1235	0.31	683	0.34	552	0.28
rs1060915 genotype frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Genotype	Count	Proportion	Count	Proportion	Count	Proportion
C/C	152	0.08	107	0.11	45	0.04
T/C	931	0.47	469	0.47	462	0.46
T/T	917	0.46	424	0.42	493	0.49
rs1060915 association with disease risk (n=2000)						
Model	**Genotype**	**Controls**	**Cases**	**OR (95% CI)**	**P-value**	
Co-dominant	T/T	424 (42.4%)	493 (49.3%)	1	<0.0001	
	T/C	469 (46.9%)	462 (46.2%)	0.85 (0.71-1.02)		
	C/C	107 (10.7%)	45 (4.5%)	0.36 (0.25-0.52)		
Dominant	T/T	424 (42.4%)	493 (49.3%)	1	0.002	
	T/C-C/C	576 (57.6%)	507 (50.7%)	0.76 (0.63-0.90)		
Recessive	T/T-T/C	893 (89.3%)	955 (95.5%)	1	<0.0001	
	C/C	107 (10.7%)	45 (4.5%)	0.39 (0.27-0.56)		
Over-dominant	T/T-C/C	531 (53.1%)	538 (53.8%)	1	0.75	
	T/C	469 (46.9%)	462 (46.2%)	0.97 (0.82-1.16)		
Log-additive	—	—	—	0.71 (0.62-0.82)	<0.0001	

4.2.1.3 BRCA1 -215T<C exon-3 (g.43115511)

The exon3 -215T<C (g.43115511) nucleotide change is novel and has not been previously reported in other populations, confirmed through HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project, is synonymous in nature because of its presence in the intronic region. This sequence variant was detected in both cases and control samples suggesting that it is a polymorphism rather than a

causative mutation. It was therefore investigated for its possible association with the disease risk through unconditional logistic regression and was observed highly significantly associated with an increased disease risk in all genetic models (Table 4.6 and Figure 4.5).

TABLE 4.6: Genotype frequencies and association of T<C transition at intron 3 of BRCA1 gene with disease risk.

Allelic frequencies of -215T<C exon-3 (n=2000)						
	All subjects		**Controls**		**Cases**	
Allele	Count	Proportion	Count	Proportion	Count	Proportion
T	3592	0.9	1898	0.95	1694	0.85
C	408	0.1	102	0.05	306	0.15
Genotype frequencies of -215T<C exon-3 (n=2000)						
	All subjects		**Controls**		**Cases**	
Genotype	Count	Proportion	Count	Proportion	Count	Proportion
C/C	112	0.06	33	0.03	79	0.08
T/C	184	0.09	36	0.04	148	0.15
T/T	1704	0.85	931	0.93	773	0.77
Association of -215T<C exon-3 with disease risk (n=2000)						
Model	**Genotype**	**Controls**	**Cases**	**OR (95% CI)**	**P-value**	
Co-dominant	T/T	931 (93.1%)	773 (77.3%)	1	<0.0001	
	T/C	36 (3.6%)	148 (14.8%)	4.95 (3.40-7.22)		
	C/C	33 (3.3%)	79 (7.9%)	2.88 (1.90-4.38)		
Dominant	T/T	931 (93.1%)	773 (77.3%)	1	<0.0001	
	T/C-C/C	69 (6.9%)	227 (22.7%)	3.96 (2.98-5.27)		
Recessive	T/T-T/C	967 (96.7%)	921 (92.1%)	1	<0.0001	
	C/C	33 (3.3%)	79 (7.9%)	2.51 (1.66-3.81)		
Over-dominant	T/T-C/C	964 (96.4%)	852 (85.2%)	1	<0.0001	
	T/C	36 (3.6%)	148 (14.8%)	4.65 (3.19-6.77)		
Log-additive	—	—	—	2.31 (1.89-2.82)	<0.0001	

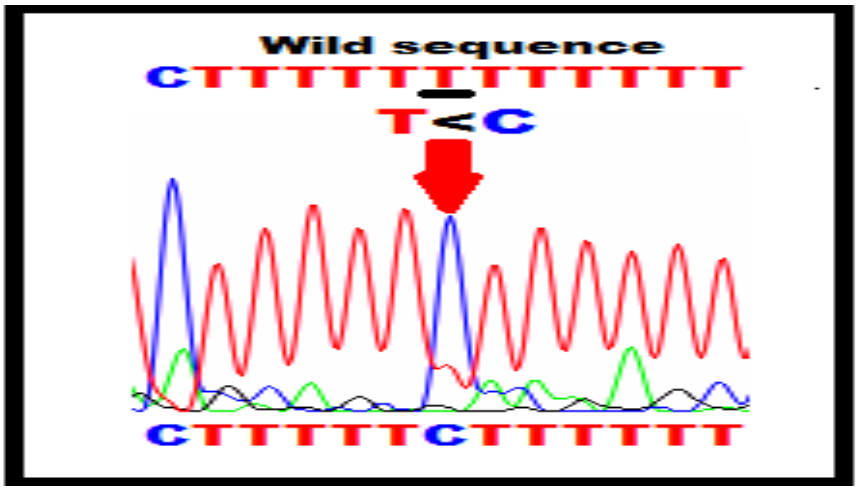

FIGURE 4.5: Direct sequencing results of BRCA1 exon-3 -215T<C transition at g.43115511, red arrow in figure points at the mutation.

4.2.1.4 BRCA1 exon-14, 102-103insTC (g.43074419-43074420)

This insertion mutation at position 102-103 at exon-14 (g.43074419-43074420) has also not previously been reported in any other population and thus is novel, confirmed through HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project. The mutation is silent at the site of insertion (ATT < ATC (I<I)), but due to frame shift the protein sequence gets changed from residues number 1529 onwards and introduces a stop codon at position 1548 and thus results in a truncated and abnormal protein. The protein will thus lack the terminal 365 residues (Figure 4.6). The mutation is part of an interpro domain named Brca1-Associated (IPR031099) also annotated as part of an interpro domain named Breast Cancer Type 1 Susceptibility Protein (BRCA1) (IPR011364). The mutation was observed in homozygous condition only in 227 (22.7%) out of the 1000 breast cancer samples sequenced and was found absent in the corresponding control samples indicating that it is disease-causative mutation. It can be suggested from our results that this mutation can be considered as a genetic risk factor for breast cancer in the Pakistani population (Figure 4.6).

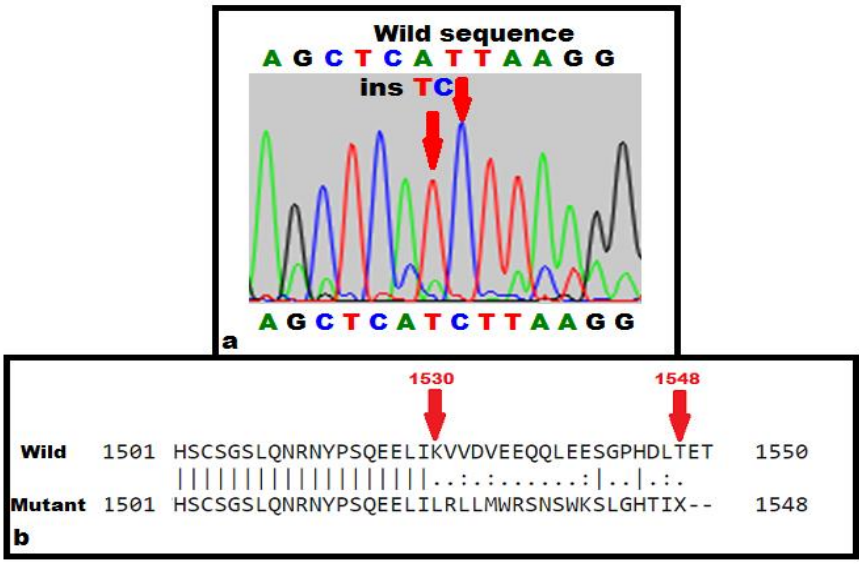

FIGURE 4.6: BRCA1 exon-14, 102-103insTC variant (a) Sequence chromatogram showing TC insertion (b) Alignment view showing change of residues due to frame shift and introduction of stop codon at position 1548 represented by red arrows.

4.2.2 Haplotype and Linkage Disequilibrium Analysis of BRCA1 Polymorphic Variants

First of all frequencies were calculated for all the haplotypes including all the 4 SNPs. As shown in Table 4.7, six common haplotypes of BRCA1 accounted for >82% of all haplotypes, the remaining were having frequency of <3% and were thus not included in the study. Haplotype analysis of the four studied polymorphic variants of BRCA1 revealed that Haplotype 3 (CATT) and 6 (AACT) were significantly associated with increased breast cancer risk (OR=15.1, 95%CI: 10.40-21.89, P<0.0001 and OR=3.24, 95%CI: 1.96-9.06, P<0.0001). In contrast, we also identified two inversely associated haplotypes (AGTT and AATT) with the disease risk (Table 4.7).

Pairwise Linkage Disequilibrium (LD) analysis of the four BRCA1 SNPs showed that the strong LD (D′ >0.52) among the 4 SNPs exists in between SNP 1 and 4

while the weakest was observed in between SNP 2 and 4. Our results of Haplotype analysis suggests that BRCA1 haplotypes and/or polymorphisms may temper breast cancer risk (Figure 4.7).

TABLE 4.7: Distribution of Haplotypes in BRCA1 Gene and their Association with breast cancer risk.

Haplotypes	**SNPs**				**Total**	**Cases freq**	**Cont freq**	**OR(95% CI)**	**P Value**
	1	**2**	**3**	**4**					
1	C	G	T	T	0.2215	0.202	0.273	1	—
2	A	G	T	T	0.1818	0.151	0.213	0.5(0.42-0.59)	0.66
3	C	A	T	T	0.1271	0.234	0.020	15.1(10.40-21.89)	<0.0001
4	C	G	C	T	0.1143	0.111	0.118	1.2(0.99-1.44)	0.58
5	A	A	T	T	0.0911	0.061	0.120	0.7(0.62-0.94)	0.11
6	A	A	C	T	0.0837	0.053	0.173	3.24 (1.96-9.06)	<0.0001

1 = rs28897696, 2 = rs28897686, 3 = rs1060915, 4 = -215 exon 3 T<C

Cases freq =Cases frequency, Cont freq = Controls frequency

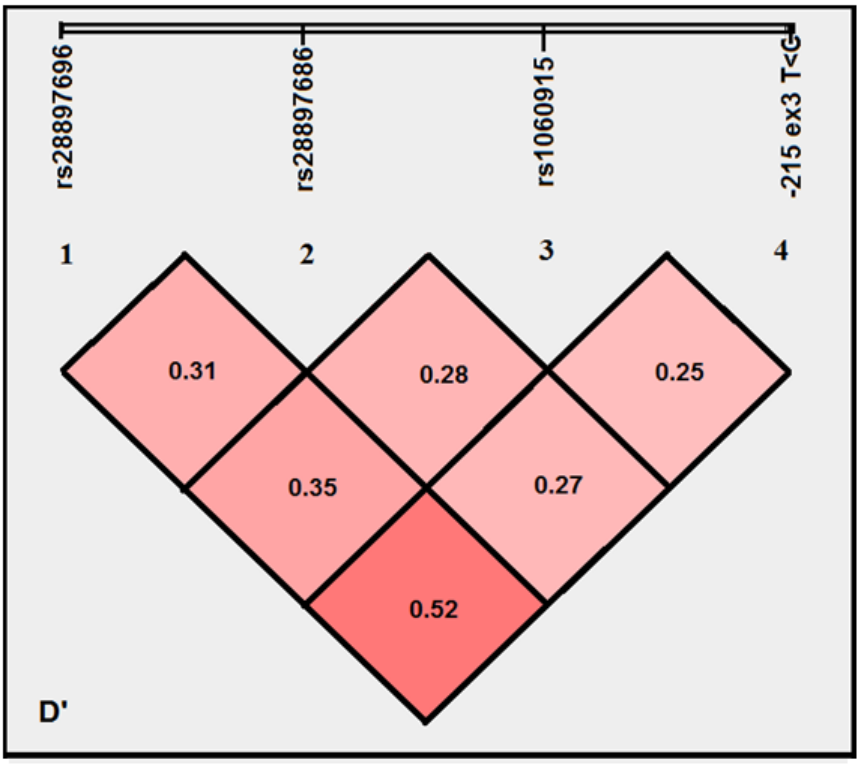

FIGURE 4.7: Analysis of Linkage disequilibrium for the investigated BRCA1 SNPs. The respective haplotype blocks are below the SNP names along with respective D' values at each intersection in the blocks.

4.2.3 BRCA2 Variants

4.2.3.1 BRCA2 Exon8 +87insA (g.32329579)

This novel insertion variant at position +87 (g.32329579) of exon-8 was confirmed as a novel from HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project. The variant was observed only homozygous condition in 19 (1.9%) out of the 1000 disease samples and was not detected in the corresponding control samples. This intronic insertion is most likely to have no impact on breast cancer development. This data suggests that the observed BRCA2 exon8 +87insA mutation is simply a result of the breast cancer phenotype rather than the selected events driving cancer development and/or progression (Figure 4.8).

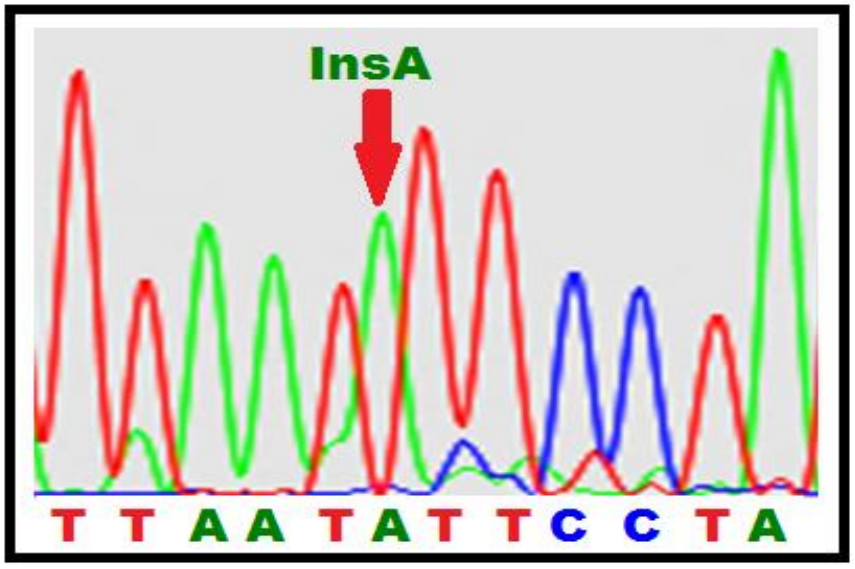

FIGURE 4.8: Direct sequencing results of BRCA2 exon-8 +87insA at g.32329579, red arrow in figure points at insertion mutation.

4.2.3.2 BRCA2 Exon-20 +318T<A (g.32371418)

The transition variant T<A at g.32371418 is novel in nature as was confirmed from HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project. Because of its presence in both cases and control, the variant was further investigated for its possible association with disease risk.

Analysis of the genotype distribution showed a significant increased risk with the +318T<A variant, with ORs equal to 3.91(95% CI=2.80-5.46) and 2.09(95% CI=1.33-3.27) for co-dominant, 3.21(95% CI=2.44-4.21) for dominant, 1.82(95% CI=1.16-2.85) for recessive, 3.78(95% CI=2.71-5.27) for over-dominant and 2.08(95% CI=1.70-2.55) for log-additive models respectively. These results suggested a possible contribution of +318T<A variant in breast cancer risk and is suggested as a genetic risk factor for tumor development in Pakistani women (Figure 4.9 and Table 4.8).

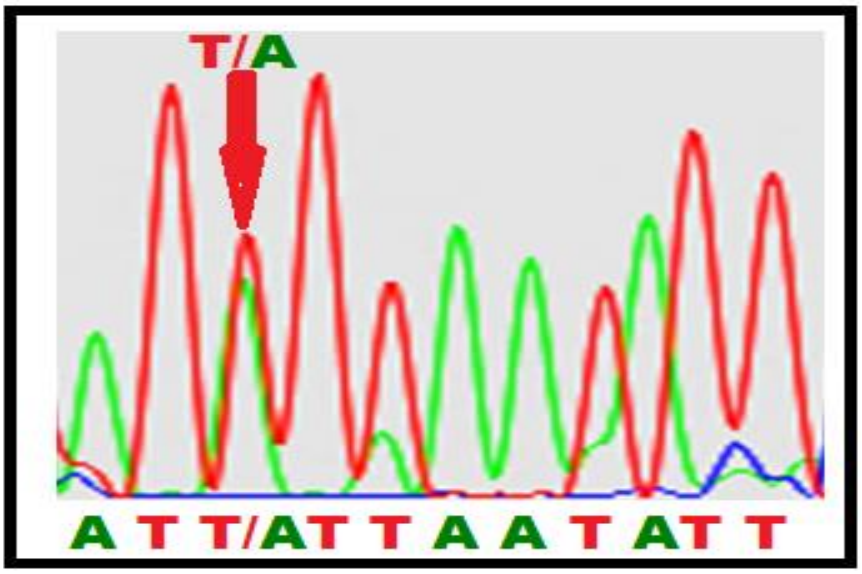

FIGURE 4.9: Sequence chromatogram showing BRCA2 Exon-20 +318T<A variant. The transition is represented by red arrow in the chromatogram.

4.2.3.3 BRCA2 exon19 -351-353delTCT (g.32370049-g.32370051)

This deletion mutation (delTCT) at BRCA2 exon19 starting at g.32370049 has not been previously reported in any world population. It was confirmed as a novel one from HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project. The variant was observed in only 283 (28.3%) out of the 1000 breast cancer samples but was not detected in any of the corresponding controls.

It can also be suggested on the basis of available data that this mutation can be considered as a genetic risk factor for breast cancer in Pakistani women (Figure 4.10).

TABLE 4.8: Frequencies and association of BRCA2 exon20 +318T<A variant with disease risk.

BRCA2 Exon-20 +318T<A allele frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Allele	Count	Proportion	Count	Proportion	Count	Proportion
T	3616	0.9	1889	0.94	1727	0.86
A	384	0.1	111	0.06	273	0.14
BRCA2 Exon-20 +318T<A genotype frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Genotype	Count	Proportion	Count	Proportion	Count	Proportion
A/A	86	0.04	31	0.03	55	0.06
T/A	212	0.11	49	0.05	163	0.16
T/T	1702	0.85	920	0.92	782	0.78
BRCA2 Exon-20 +318T<A association with disease risk (n=2000)						
Model	**Genotype**	**Controls**	**Cases**	**OR (95% CI)**	**P-value**	
Codominant	T/T	920 (92%)	782 (78.2%)	1	<0.0001	
	T/A	49 (4.9%)	163 (16.3%)	3.91 (2.80-5.46)		
	A/A	31 (3.1%)	55 (5.5%)	2.09 (1.33-3.27)		
Dominant	T/T	920 (92%)	782 (78.2%)	1	<0.0001	
	T/A-A/A	80 (8%)	218 (21.8%)	3.21 (2.44-4.21)		
Recessive	T/T-T/A	969 (96.9%)	945 (94.5%)	1	0.0078	
	A/A	31 (3.1%)	55 (5.5%)	1.82 (1.16-2.85)		
Overdominant	T/T-A/A	951 (95.1%)	837 (83.7%)	1	<0.0001	
	T/A	49 (4.9%)	163 (16.3%)	3.78 (2.71-5.27)		
Log-additive	—	—	—	2.08 (1.70-2.55)	<0.0001	

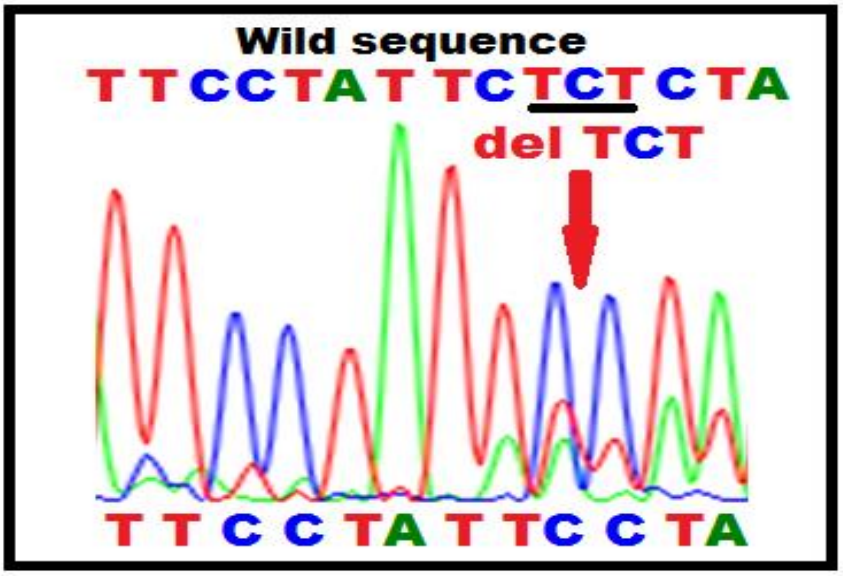

FIGURE 4.10: Identification of novel BRCA2 exon19 delTCT mutation, the red arrow marks the start of deletion.

4.2.3.4 BRCA2 exon16 -17G<T (g.32357725)

The transition variant exon16 -17G<T at position g.32357725 is novel in nature as was confirmed from HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project. The variant was detected in both cases and control samples and thus can be regarded as a polymorphism. The variant was investigated in detail for its potential association with the disease risk in case-control samples.

Results of logistic regression showed a positive association of -17G<T variant on co-dominant model (GG vs TT) with OR=1.43, 95%CI=1.02-1.99, Dominant model (OR=1.29, 95%CI=1.04-1.61) and Log-additive model (OR=1.20, 95%CI=1.03-1.39) respectively. These results suggested a possible contribution of -17G<T polymorphism in breast cancer Pakistani women (Figure 4.11 and Table 4.9).

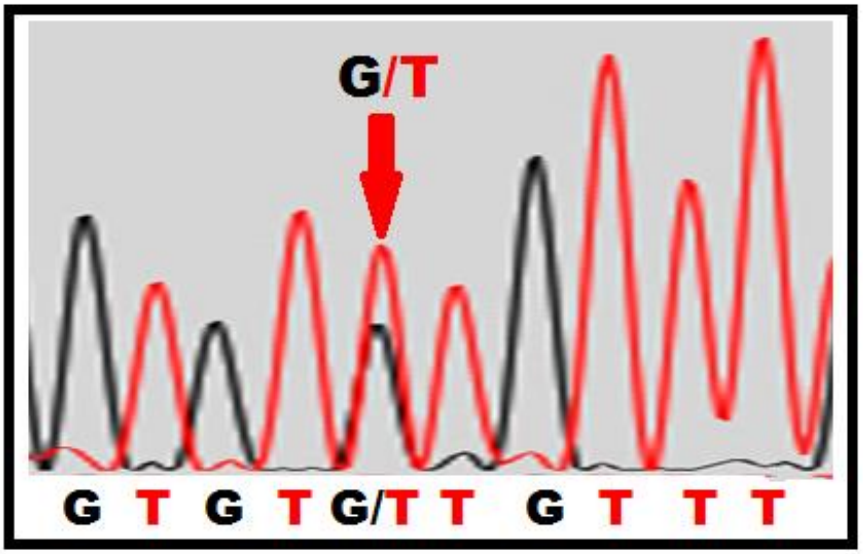

FIGURE 4.11: Direct sequencing results of BRCA2 exon16 -17G<T at g.32357725, red arrow in figure points at substitution event.

4.2.3.5 BRCA2 exon27 T129A (g.32398290)

This transversion of nucleotide T<A of BRCA2 exon27 has been confirmed from HGVD/HGMD, Ensembl, ExAC and 1000 Genomes Project that it has not been reported in any other population and thus is novel. The variant was observed in 339 (33.9%) out of the 1000 breast cancer samples only and was not found in any of the corresponding control samples. The transversion results in the substitution of Aspartic acid to Glutamic acid at position 3260. There is no change of net charge due to the mutation but the mutant and wild-type residues differ in size. The mutant residue is bigger than the wild type, this might lead to bumps in the protein structure. This residue is part of an interpro domain named: Breast Cancer Type 2 Susceptibility Protein IPR015525 (Figure 4.12).

Our data indicates that the variant is a disease causative mutation rather than a polymorphism. It can also be suggested on the basis of available data that this mutation can be considered as a genetic risk factor for breast cancer in Pakistani women. The variant was functionally predicted as Low (1.93) by MutationAssessor, probably benign by PANTHER, Neutral (-0.875) by PROVEAN, Tolerated (2.0) by SIFT and Benign (0.208) by PolyPhen2.

TABLE 4.9: Frequencies and association of BRCA2 exon16 -17G<T variant with disease risk.

Allele frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Allele	Count	Proportion	Count	Proportion	Count	Proportion
G	3444	0.86	1754	0.88	1690	0.84
T	556	0.14	246	0.12	310	0.16
Genotype frequencies (n=2000)						
	All subjects		**Controls**		**Cases**	
Genotype	Count	Proportion	Count	Proportion	Count	Proportion
G/G	1597	0.8	819	0.82	778	0.78
G/T	250	0.12	116	0.12	134	0.13
T/T	153	0.08	65	0.06	88	0.09

Association of BRCA2 -17G<T with disease risk (n=2000)					
Model	**Genotype**	**Controls**	**Cases**	**OR (95% CI)**	**P-value**
Co-dominant	G/G	819 (81.9%)	778 (77.8%)	1	0.054
	G/T	116 (11.6%)	134 (13.4%)	1.22 (0.93-1.59)	
	T/T	65 (6.5%)	88 (8.8%)	1.43 (1.02-1.99)	
Dominant	G/G	819 (81.9%)	778 (77.8%)	1	0.022
	G/T-T/T	181 (18.1%)	222 (22.2%)	1.29 (1.04-1.61)	
Recessive	G/G-G/T	935 (93.5%)	912 (91.2%)	1	0.053
	T/T	65 (6.5%)	88 (8.8%)	1.39 (0.99-1.94)	
Over-dominant	G/G-T/T	884 (88.4%)	866 (86.6%)	1	0.22
	G/T	116 (11.6%)	134 (13.4%)	1.18 (0.90-1.54)	
Log-additive	—	—	—	1.20 (1.03-1.39)	0.016

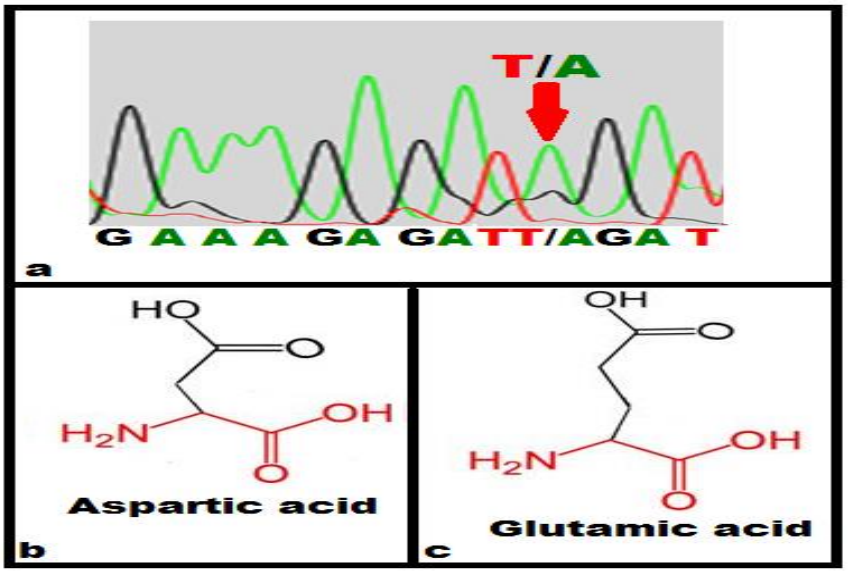

FIGURE 4.12: BRCA2 exon27 T129A transversion at g.32398290 (a) Direct sequencing results of BRCA2 exon27 T129A at g.32398290, red arrow in figure points at the substitution event (b) Wild residue aspartic acid (c) Mutant residue Glutamic acid.

4.2.4 Haplotype and Linkage Disequilibrium Analysis of BRCA2 Gene

Haplotype frequencies were calculated for all the haplotypes including the observed four SNPs of BRCA2. Haplotype 1 was the most commonly observed haplotype having a frequency of 0.77 and was thus set as a reference haplotype. Five common haplotypes of BRCA1 as shown in Table 4.10 accounted for approximately 99% of all haplotypes, the remaining rare haplotypes were thus ignored during the study. The analysis revealed three positively associated Haplotypes (CGAG, CGAT and CATG) with breast cancer risk (Table 4.10).

Pairwise Linkage Disequilibrium (LD) analysis of the four BRCA2 SNPs showed that the strong LD (D'=0.43) exists in between SNP 2 and 4 followed by (D'=0.35) SNP1 and 3 (Figure 4.13).

TABLE 4.10: Distribution of Haplotypes in BRCA2 Gene and their Association with breast cancer risk.

Haplotypes	**SNPs**				**Total**	**Cases freq**	**Cont freq**	**OR(95% CI)**	**P Value**
	1	**2**	**3**	**4**					
1	C	G	T	G	0.773	0.724	0.829	1.00	—
2	C	G	T	T	0.115	0.113	0.11	1.14(0.97-1.35)	0.12
3	C	G	A	G	0.070	0.091	0.041	2.03(1.60-2.57)	<0.0001
4	C	G	A	T	0.022	0.032	0.012	2.38(1.50-3.77)	<0.0001
5	C	A	T	G	0.013	0.021	0.004	4.03(1.95-8.34)	<0.0001

1 = rs4987049, 2 = rs28897743, 3 = +318T¿A, 4 = -17G<T

Cases freq = Cases frequency, Cont freq = Controls frequency

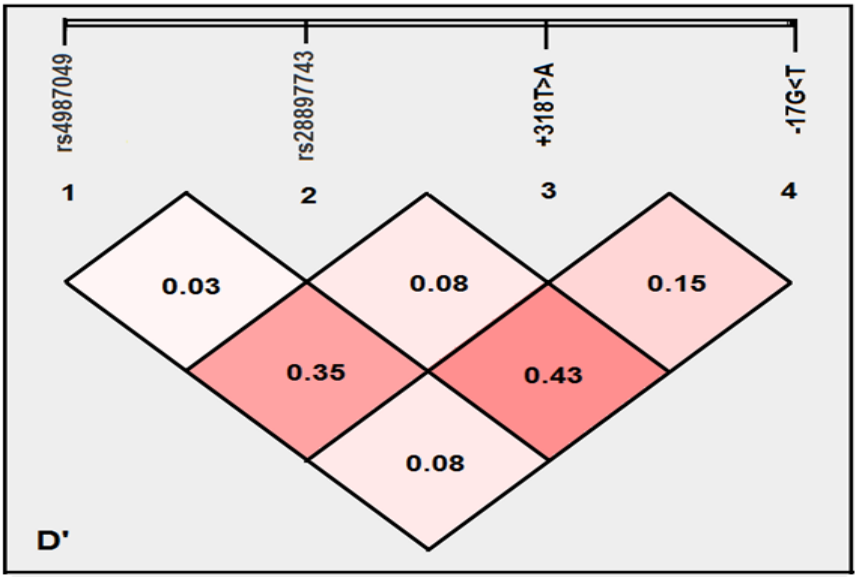

FIGURE 4.13: Analysis of Linkage disequilibrium for the investigated BRCA2 SNPs. The respective haplotype blocks are below the SNP names along with respective D' values at each intersection in the blocks.

BRCA1 has 24 exons, covering a length of about 100kb of genomic DNA and encodes a protein of 1863 residues. The RING finger domain of the protein (amino acids 20-64) is at the NH2 terminus and exon-11 comprises of the interaction domains with RAD51 [157], RAD50 complex [158] and FANCA [159] etc. The BRCT domain lies in between residues 1646-1863 is a conserved domain in BRCA1

[160]. The BRCT domain binds to phosphorylated partner proteins by recognizing phosphopeptide and is involved in the DNA damage response [161], [162]. The C-terminal region is a transcriptional co-regulator and has its function in cell cycle control response [160], [163], [164].

BRCA2 is larger than BRCA1 gene, consisting of 27 exons and spans about 70kb. It encodes a protein of 3418 residues. Eight BRC repeats locating in exon-11 comprise of the Rad51 binding domain [165]. The C-terminal domain is well conserved [166] and binds to ssDNA and DSS1 (deleted in split-hand/split-foot syndrome 1) [167], which is required for homologous-recombination.

BRCA1 and BRCA2 proteins have various functions including DNA repair by homologous recombination, regulation of cell cycle checkpoints and transcriptional regulation [168]. BRCA1 and BRCA2 have a direct part in DNA repair. BRCA2 and RAD51 communicates and co-localize in a BRCA1-BRCA2-RAD51 complex [157], [169]. Moreover, BRCA1 also interacts with another protein RAD50, which is involved in homologous recombination and DNA damage response [158]. BRCA1 and BRCA2 both have their role repair of DNA double-strand breaks (DSBs) through homologous and non-homologous recombination [157], [158].

BRCA1 and BRCA2 genes are essential for preserving chromosome structure and thus gross chromosomal rearrangements (GCRs) were demonstrated by human cancer cells lacking both or any of these genes [170], [171]. It can thus be concluded that these genes may play a caretaker role in preserving genomic integrity. Another function of BRCA1 is in checkpoint control through its BRCT motif situated at its carboxyl terminus [172]. BRCA1 protein is also associated with the centrosome amplification and the G2/M transition [173], [174].

In this study, we focused on identifying BRCA1 and BRCA2 genetic alterations among Pakistani women affected with Breast Cancer. We have detected a total of thirteen variants in BRCA1 and BRCA2 respectively including three novel variants in BRCA1 and five novels in BRCA2. Along with the novel genetic alterations, three previously reported SNPs in BRCA1 (rs28897686, rs28897696 and rs1060915)

and two in BRCA2 (rs4987049 and rs28897743) have also been reported in this study.

The rs28897686 (E1250K) SNP of BRCA1 substitutes Glutamic acid to Lysine at position 1250. This SNP was observed positively associated with the disease risk and is thus a risk factor for Pakistani women. rs28897696 (A1708E) changes Alanine to Glutamic acid but was found not associated with the breast cancer in study population. The third SNP of BRCA1, rs1060915 (S1436S) is synonymous in nature and thus is silent. This SNP was found associated with health according to our results. The novel variants of BRCA1 -37insC of exon3 (Figure 4.3) and -215T<C of exon3 (Figure 4.5) are intronic in nature having no impact on protein structure. They were found associated with the disease risk and were suggested as risk factors for breast cancer in Pakistani population. BRCA1 102-103insTC (Figure 4.6) in exon-14 is exonic in nature and was detected only breast cancer cases.

The rs4987049 (Y3308X) SNP introduces a stop codon at position 3308 in BRCA2 and thus results in a truncated protein while rs28897743 (R2336H) substitutes Arginine to Histidine and was found significantly associated with breast cancer risk. The novel variants, exon8 +87insA (Figure 4.8), exon20 +318T<A (Figure 4.9), exon19 -351-353delTCT (Figure 4.10) and exon16 -17G<T (Figure 4.11) are intronic variants having no impact on protein structure but were found associated with the disease risk. T129A mutation is the only reported novel variant of BRCA2 in exon 27 in this study which results in substitution of Aspartic acid to Glutamic acid at position 3260.

The majority of BRCA1 mutations cause protein truncations due to the introduction of stop codons due to small mutations [175]. Large deletions are rarely been reported in BRCA1, comprising 5-10% of all germ line mutations. They are even less common in BRCA2 [176]. In rare cases, complex rearrangements in BRCA1 including Alu sequences were reported [177]. Over 300 missense mutations of BRCA1 have already been reported in different databases. The functional impact of these mutations is still not well understood and is considered as a challenging

problem for researchers to evaluate the risk of having cancer in women carrying any of this missense mutation.

Some BRCA1/2 mutations have already been profiled in specific ethnic populations. The most prevalent founder mutations of BRCA1 and BRCA2 were identified by Simard and colleagues [178]. They had identified BRCA1 185delAG and 5382insC as recurrent mutations from Quebec, Canada with strong Ashkenazi ancestry. Offit and Neuhausen identified 6174delT mutation as founder mutation of BRCA2 with Ashkenazi Ancestry [179]. Founder mutations have also been identified in Austrian [180], Icelandic, Dutch, Norwegian, French Canadian [181], Belgian [182], British [183], Finnish [184] and Swedish [185] populations.

4.3 Socio-demographic and Reproductive Risk Factors

Though every women is at risk of having breast cancer regardless of their ethnic and racial backgrounds but there exists a differential variation in incidence rate among different populations. It suggests that etiological variables might differ in their biologic expression and thus affecting the disease onset [186]. Population based cancer registries are lacking in the developing countries like Pakistan and most of the available figures are centered on data from small units of the population [57]. It is therefore vital to explore the likely factors for in Pakistani women, which might contribute to current knowledge of this vital topic. The primary aim of this project was to examine the cross-sectional associations of life style, reproductive and socio-demographic risk factors with breast cancer density in Pakistani women.

Mean (±standard deviation) age of cases and controls at recruitment was 50.58±10.68 and 54.78±14.52 years respectively. Cases were comparatively younger than the controls. BMI was calculated for all the cases and controls by recording their weight (kg) and height (m2). Mean BMI for BC cases and controls were 26.07±4.04 and 25.05±4.25 respectively while range of BMI for both cases and controls was

observed as 15.32-35.20 and 17.0-35.40. Among the patients 60.70% were married, 46.50% were nulliparous, 16.90% had 4 or >4 children, 39.90% females breast fed their children. Considering smoking and physical activity, 88.90% were non-smokers and 67.90% were physically active. Post-menopausal women diagnosed with breast cancer accounted for 52.30% (Table 4.11).

TABLE 4.11: Clinical and Socio-demographic characteristics and potential Breast cancer risk factors in Pakistani Population.

Factors	**Cases (1000)**		**Controls (1000)**	
	Number	**Percentage**	**Number**	**Percentage**
Age (years)				
<40	139	13.90	100	10.00
40-49	362	36.20	316	31.60
50-59	330	33.00	285	28.50
60-69	102	10.20	74	7.40
70-79	50	5.00	158	15.80
80-89	17	1.70	67	6.70
BMI				
<18.5	11	1.11	91	9.10
18.5-25	343	34.30	405	40.50
25-30	438	43.80	339	33.90
≥30	217	21.70	165	16.50
Marital Status				
Single	393	39.30	242	24.20
Married	607	60.70	758	75.80
Parity (number of children)				
0	465	46.50	326	32.60
1	107	10.70	70	7.00
2	138	13.80	75	7.50
3	121	12.10	195	19.50
4+	169	16.90	334	33.40
Oral contraceptive use				

Factors	Cases (1000)		Controls (1000)	
	Number	**Percentage**	**Number**	**Percentage**
Yes	691	69.10	397	39.70
No	309	30.90	603	60.30
Breast Feeding				
Yes	399	39.90	636	63.60
No	601	60.10	364	36.40
Smoking				
Yes	111	11.10	74	7.40
No	889	88.90	926	92.60
Physical Activity				
Yes	679	67.90	788	78.80
No	321	32.10	212	21.20
Menopause status				
Pre-menopause	477	47.70	551	55.10
Post-menopause	523	52.30	449	44.90

Table 4.12 represents the basic clinical features for breast cancer cases. 58.10% were the females diagnosed with right breast affected, 33.20% with left breast and 8.70% with both breast affected. Among the breast cancer cases, 67.40% were being under treatment with chemotherapy and 17.40% were treated with mastectomy.

Breast cancer can be divided into three basic stages on the basis of cancer cells, i) Local disease, where cancerous cells have yet not been spread external to the breast or only a small proportion of cancer cells are within the lymph nodes close to the breast, ii) Locally advanced breast cancer is the type in which cancer cells can be detected in armpits or in the lymph nodes near the breast-bone or small areas of cancer cells are in the lymph nodes and iii) Distant metastasis refers to cancer which has been spread to other parts of the body. Disease stage was

recorded local in 58.70%, locally advanced in 39.40% and 1.90% were diagnosed with distant metastasis.

Some people are comparatively at higher risk of developing the disease due to various genetic factors. Family history, especially having a close blood relative diagnosed with cancer, doubles the risk of breast cancer. In the current data set, 31.70% patients had at least a blood relative diagnosed with some type of cancer, mostly breast cancer. It was also observed that 22.80% patients were diagnosed with other types of medical complications including high blood pressure, diabetes etc. as well.

Our results revealed significant association between different factors recorded for breast cancer patients. There was a significant association between age and breast cancer. Females having ages from 40 to 69 years, patients aged in 40s (OR = 1.22, 95% CI: 1.03-1.46), 50s (OR = 1.24, 95% CI: 1.03-1.49) and patients aged 60s (OR = 1.41, 95% CI: 1.05-1.90) were more likely to have breast cancer.

TABLE 4.12: Distribution of cases by Pathologic features of breast cancer patients in Pakistani population.

Patients pathologic characteristics	**Number**	**Percentage**
Breast Affected		
Right Breast	581	58.10
Left Breast	332	33.20
Bilateral	87	8.70
Medical Treatment		
Chemotherapy	674	67.40
Radiotherapy	186	18.60
Both	140	14.00
Surgery Status		
Mastectomy	174	17.40
Lumpectomy	324	32.40
No	502	50.20

Patients pathologic characteristics	**Number**	**Percentage**
Disease Stage		
Local	587	58.70
Locally Advanced	394	39.40
Distant	19	1.90
Other Complications		
Cancer + (any other including Diabetes, Diabetes + HBP, HBP or Others)	228	22.80
No	772	77.20
Family History		
Yes	317	31.70
No	683	68.30

Obesity has a vital role in breast cancer growth and development [187]. To examine the influence of this significant risk factor on Pakistani women with breast cancer, we examined this link among breast cancer cases and controls. An increase in the basic metabolic index is positively correlated with breast cancer and thus the proportion of females having BMI≥25 was significantly higher among the patients. Overweight (BMI≥25) and obese (BMI≥30) females have approximately 1.5 times more risk of having breast cancer (Overweight; OR=1.57, 95% CI: 1.28-1.93 and Obese; OR=1.60, 95% CI: 1.26-2.03). So undoubtedly, obesity is one of the contributor for enhancing breast cancer risk among Pakistani women.

According to reports, 1 out of nine women in Pakistan is likely to have breast cancer. These figures indicates one of the highest rate of breast cancer incidence in Asia. The incidence rate in Pakistani women has been reported 50 out of one lakh which is much higher than the neighboring countries [188]. In 2013-2014, an observational study was carried out based on breast cancer patients at Shaukat Khanum Memorial Cancer Hospital and Research Center, Pakistan. According to the results of this study 4,366 breast cancer cases were recorded including approximately 80.4% of the cases belonged to the province of Punjab [189]. According

to Pakistan Atomic Energy Commission Cancer Registry the incidence of breast cancer in Pakistani women is 46.7% based on the data collected from different nuclear medicine and oncology institutes for 30 years (1984-2014) [190].

It was also observed that unmarried women were at more than two fold higher risk (OR=2.03, 95% CI: 1.69-2.44). Nulliparous women had higher risk for breast cancer (OR=2.56, 95% CI: 1.87-3.51) compared to parous women. It was further observed that the risk decreases with increase in parity. Among the parous women never breastfeeding females were have more than 1.5 fold higher risk of developing breast cancer (OR=1.82, 95% CI: 1.35-2.46). Similarly use of oral contraceptives (OR=3.41, 95% CI: 2.86-4.06), and smoking (OR=1.56, 95% CI: 1.16-2.09) were also significantly associated with increasing risk of breast cancer. Individuals who were physically inactive were recorded to be 1.27 times more likely to develop breast cancer than those who are physically active (OR=1.75, 95% CI: 1.44-2.12). When the menopausal status was studied, we have found approximately 1.34 fold increase in the disease risk among the postmenopausal patients (OR =1.34, 95% CI: 1.14-1.58) (Figure 4.14). A total of 269 patients were censored during the study out of the total 1000. The patients had an overall median survival time of 33 months (95% CI: 28-34) (Figure 4.15).

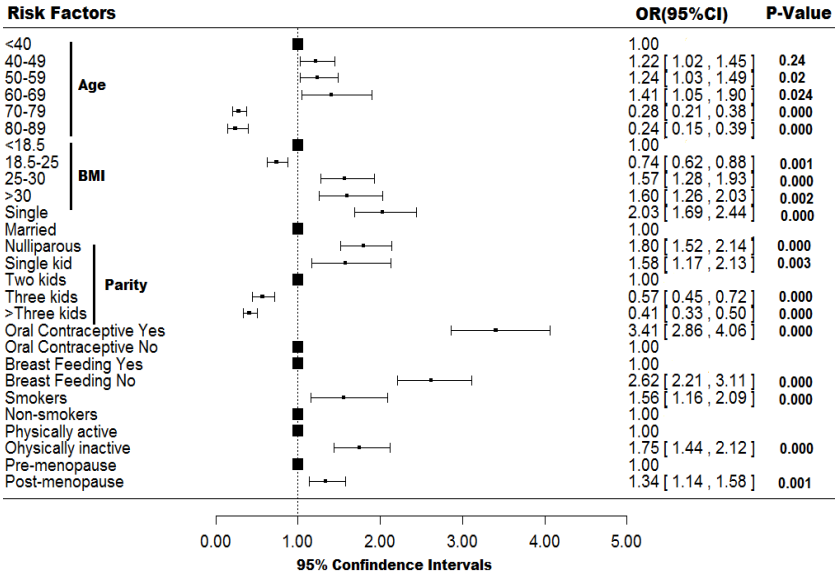

FIGURE 4.14: Association of risk factors with breast cancer in the study population.

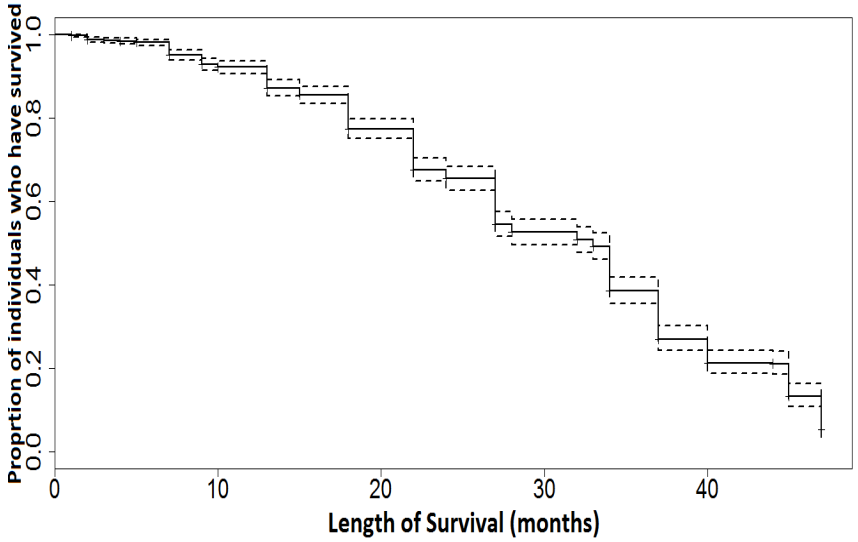

FIGURE 4.15: Overall survival curve (Kaplan-Meier plot) of Breast Cancer Patients. Solid line represents survival while dashed lines represent 95% confidence intervals (95% CI).

Breast cancer is the most frequently diagnosed cancer among Pakistani females [53] and its incidence is still increasing [80]. This ongoing rising trend in breast cancer cases has created an imperative need to develop preventive approaches. Age is an important risk factor for different cancers [6]. Higher proportion of breast cancer cases has been detected by the age of <40, 40s and 50s. Ahmadian and coworkers observed that ladies usually below the age of 40 had been affected with breast cancer as compared to rest of the world where females above 40 years are influenced [191]. Mean age at diagnosis was observed 51.40±10.40 in the present study, ranging from 29-87 years. Almost similar pattern of age at diagnosis has been reported by other researchers [192]–[194]. A study conducted on Iranian women reported that patients were mostly aged above 44 [195]. The average age breast cancer incidence among white females in the united states is 61 years [196], while it is 51.4 years in Pakistani females according to our study. It is almost a decade earlier, as compared to other parts of the world specially the Western countries. The reason for this early age incidence among Pakistani women needs to be further investigated.

Obesity is a common health problem and its frequency is globally increasing. It is well established that increased BMI is related to enhanced incidence of certain

diseases including cancers and is also associated to higher morbidity [197]. Patients having a normal BMI have a comparatively longer overall survival rates as compared to the overweight or obese patients [198]. In the current report the mean BMI was observed to be 22.64±3.19 ranging from 15.32-35.20. It was observed in this case-control study that, increasing BMI is responsible for elevating breast cancer risk. Only 33.36% of the study subjects were having BMI in the normal range (18.5-23). The risk of carcinoma increased as the BMI increased from the normal range. Other studies also have reported similar result [6], [35], [199]–[201], which may be mainly due to the fact that excess aromatase activity in the peripheral adipose tissues produced higher levels of free estrogen [6].

Tobacco use and physical idleness are the main sources of major non-communicable maladies including certain types of cancers. Both of these factors are though preventable but contribute significantly to the over-all disease burden, disability and mortality [202]. Tobacco smoking is one of leading avoidable risk variable for cancer in general [203] and its relationship with breast cancer risk is also uncertain. It is now evident that over-weight/obesity and physical inactivity are major risk factors for developing breast cancer in many countries [204]. More commonly, people in the urban areas are physically inactive and appropriate measures are therefore needed to avoid weight gain, which would be cost effective than the treatment and allied obstacles. These factors are modifiable and can be prevented.

There exists a positive association between being physically inactivity and breast cancer risk in Pakistani women (OR=1.72, 95% CI: 1.44-2.12). The relationship between breast cancer risk and physical activity is complex and therefore need to be described through molecular mechanisms from different angles. Physical activity has a multi-dimensional impact on breast cancer because it also other associated risk factors including BMI, menstrual cycle, hormones and immune system [205]. It is now well established that physical inactivity results in an increase in obesity, which has independently been considered a risk factor for breast cancer. There exists an inverse correlation between physical activity and obesity [206].

Research has showed that marital status somehow affect an individual's health but this association has not been studied comprehensively [207], [208]. Married women were less likely to have breast cancer according to our results (OR=0.49). Some other researchers have also reported an association between marital status and multiple cancers, which support our results that unmarried individuals are at greater risk in the population. Aizer and collegues found that unmarried individuals have significantly higher risk of metastatic cancer [209].

In general, reproductive factors, like parity and breastfeeding have already been revealed to have a protective significance against breast cancer [210]. Parous females have a comparatively lower risk, but this relationship is very complex [211]. In our study we observed that, breast cancer risk gets decreases with increasing parity. Higher the number of full-term pregnancies, the greater the protection. A single full-term pregnancy is able to reduce the risk of breast cancer by approximately 25% [212] and the risk further decreases up to 50% with five or more children [211], [213]. Breast feeding also has a shielding outcome against breast cancer. It was observed in the results that cases were less likely to have breastfeed (OR=0.55). These findings are consistent with the findings of some other studies [8], [214]–[216] but further research is recommended to explore the causal mechanisms that how breastfeeding influence breast cancer.

Usage of oral contraceptives has modernized the reproductive life of women [217] but only few authors have focused to ascertain the possible link of oral contraceptives with breast cancer risk. The reported studies have suggested little or no link for this association [218]. We recorded in our study a positive association between breast cancer risk and usage of oral contraceptives. Our results are in line with other studies, who also have reported that women having used oral contraceptives are at comparatively higher risk of having breast cancer [219], [220]. The hormonal effect posed by the oral contraceptives is complex. They may cause protective anovulation on one hand or may also can stimulate mitotic activity through the mixture of estrogen and progesterone in breast tissues [221]. Experimental data also support the fact that, estrogens are associated with the growth

and development of breast cancer and exert both direct and indirect proliferative effects on human breast-cancer cells [222].

Menopause is not directly related to cancer, but actually the risk of developing cancer increases with the increasing age [223]. During the reproductive age of females, the ovaries produce steroid hormones affecting function and development of the breast [224].

In conclusion, the current data support that various risk factors including age, BMI, marital status, parity, oral contraceptives, breast-feeding, smoking, physical activity and menopausal status were significantly associated with increased risk of developing breast cancer in Pakistani women. The role of breastfeeding, ages at menarche and menopause needs clarification and further work. It would be useful to confirm these findings in additional studies that include area-based data to capture the ethnic differences in breast cancer cases.

4.4 Body Mass Index (BMI) and Risk of Breast Cancer in Pakistani Women

Overweight and obesity are foremost communal health concerns in both developing and developed countries [225], which increases the risk of several chronic diseases including cancers [226], [227]. According to an estimate, about 20% of cancers are caused due to excess weight gain [228]. The Million Women Study, which is the biggest one on ladies has demonstrated that roughly half can be credited to obesity in postmenopausal women [229]. A number of studies examining the relationship between BMI and malignancy suggested that a higher BMI can increase tumor rate. Maintaining a healthy weight has been proposed in tumor avoidance procedures [230], recommending a conceivable role of changes in BMI and also baseline BMI with respect to cancer risk. The mechanisms underlying the basic connection between BMI and cancer risk are still inadequately understood [227].

High BMI has also been documented as an important and positively associated risk factor for breast cancer in many other previous epidemiological studies [231]–[234]. It is unfortunate that majority of these studies have been carried out in Western countries. There exists clear cut variation in the frequency and mortality of BCa in different geographical regions which suggests that known factors may vary in different parts of the world. The frequency of breast cancer has dramatically augmented in Pakistani women which are predicted mainly due to adaptation of westernized life style, which ultimately leads to an increased BMI. This part of study was thus aimed to examine the association between BMI and BCa risk among Pakistani ladies.

Majority of study subjects were having their BMI >25kg/m^2. About 34.41% of women were in the BMI range <25km/m^2, 43.86% were in 25-29.9 (over weight) and 21.73% were ≥30kg/m^2 (obese). There were significant differences between cases and controls for all the studied risk factors. The percent distribution of certain breast cancer risk factors varied considerably by BMI (Table 4.13).

TABLE 4.13: Characteristics of Pakistani breast cancer patients (women) according to Body Mass Index (BMI).

Features	**Body Mass Index (Kg/m^2)**					
	<25		**25-29.9**		**≥30**	
	Cases (426)	**Controls (500)**	**Cases (543)**	**Controls (342)**	**Cases (269)**	**Controls (166)**
Age	n(%)	n(%)	n(%)	n(%)	n(%)	(%)
<40	31(7.3)	66(13.2)	126(23.2)	34(9.9)	15(5.6)	1(0.6)
40-49	144(33.8)	142(28.4)	146(26.9)	84(24.6)	158(58.7)	93(56.0)
50-59	137(32.2)	123(24.6)	177(32.6)	119(34.8)	95(5.6)	45(27.1)
60-69	56(13.1)	35(7.0)	69(12.7)	23(6.7)	1(58.7)	17(10.2)
70-79	39(9.2)	104(20.8)	23(4.2)	47(13.7)	0(0.0)	8(4.8)
>79	19(4.5)	30(6.0)	2(0.4)	35(10.2)	0(0.0)	2(1.2)
Marital status						
Single	181(42.5)	133(26.6)	233(42.9)	74(21.6)	73(27.1)	37(22.3)

Features	**Body Mass Index (Kg/m^2)**					
	<25		**25-29.9**		**≥30**	
	Cases (426)	**Controls (500)**	**Cases (543)**	**Controls (342)**	**Cases (269)**	**Controls (166)**
Married	245(57.5)	367(73.4)	310(57.1)	26878.4)	196(72.9)	129(77.7)
Number of kids						
0	203(47.7)	170(34.0)	277(51.0)	106(31.0)	96(35.7)	53(31.9)
1	48(11.3)	36(7.2)	58(10.7)	21(6.1)	26(9.7))	14(8.4)
2	50(11.7)	38(7.6)	76(14.0)	23(6.7)	45(16.7)	15(9.0)
3	54(12.7)	101(20.2)	54(9.9)	65(19.0)	42(15.6)	31(18.7)
>3	71(16.7)	155(31.0)	78(14.4)	127(37.1)	60(22.3)	53(31.9)
Breast Feeding						
Yes	162(38.0)	310(62.0)	199(36.6)	226(66.1)	134(49.8)	105(63.3)
No	264(62.0)	190(38.0)	344(63.4)	116(33.9)	135(50.2)	61(36.7)
Smoking						
Yes1	45(10.6)	37(7.4)	66(12.2)	25(7.3)	27(10.0)	13(7.8)
No1	381(89.4)	463(92.6)	477(87.8)	317(92.7)	242(90.0)	153(92.2)
Physical activity						
Yes2	278(65.3)	385(77.0)	349(64.3)	294(86.0)	214(79.6)	115(69.3)
No2	148(34.7)	115(23.0)	194(35.7)	48(14.0)	55(20.4)	51(30.7)
Menopausal status						
pre	248(58.2)	333(66.6)	253(46.6)	146(42.7)	90(33.5	76(45.8)
post	178(41.8)	167(33.4)	290(53.4)	196(57.3)	179(66.5)	90(54.2)

BMI category of 25-29.9 Kg/m^2 was observed associated with breast cancer risk in ages range from 40-49 (OR, 1.47; 95% CI, 1.11-1.95) and 60-69 (OR, 2.53; 95% CI, 1.57-4.08), while BMI ≥30 was found associated in ages 40-49 (OR, 1.44; 95% CI, 1.10-1.89) and 50-59 (OR, 1.78; 95% CI, 1.23-2.56). We found no evidence of increased risk of breast cancer in BMI <25Kg/m^2 in all age ranges. Further analysis of data was performed, and the results indicated that the BMI has a positive association with postmenopausal breast cancer among Pakistani women.

Overall, breast cancer risk was increased by increasing BMI especially for women with BMI 25-29.9 kg/m^2 (OR, 1.27; 95% CI, 1.03-1.55) and BMI ≥30 (OR, 1.72; 95% CI, 1.32-2.25) (Table 4.14).

TABLE 4.14: Association of different risk factors with breast cancer in Pakistani (women) according to Body Mass Index (BMI).

Characteristics	**BMI (Kg/m^2)**					
	<25		**25-29.9**		**≥30**	
Age	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value
<40	1		1		1	
40-49	0.80 (0.63-1.03)	0.08	1.47 (1.11-1.95)	0.01	1.44 (1.10-1.89)	0.01
50-59	0.90 (0.69-1.16)	0.40	1.25 (0.97-1.60)	0.08	1.78 (1.23-2.56)	0.00
60-69	1.32 (0.86-2.03)	0.21	2.53 (1.57-4.08)	0.00	0.05 (0.01-0.35)	0.00
70-79	0.28 (0.19-0.41)	0.00	0.39 (0.23-0.64)	0.00	0.05 (0.00-0.82)	0.04
>79	0.51 (0.28-0.91)	0.02	0.04 (0.01-0.19)	0.00	0.16 (0.01-3.39)	0.24
Marital status						
Single	1.13 (0.89-1.43)	0.33	2.93 (2.22-3.86)	0.00	1.64 (1.10-2.46)	0.02
Married	1		1		1	
No of kids						
0	0.97 (0.77-1.21)	0.77	2.45 (1.93-3.12)	0.00	1.51 (1.07-2.14)	0.02
1	1.09 (0.70-1.69)	0.70	2.31 (1.39-3.83)	0.00	1.52 (0.79-2.93)	0.21
2	1.07 (0.70-1.65)	0.74	2.80 (1.74-4.50)	0.00	2.50 (1.38-4.51)	0.00
3	0.41 (0.29-0.58)	0.00	0.66 (0.46-0.96)	0.03	1.11 (0.69-1.77)	0.67
>3	1		1		1	
Breast feeding						
Yes	1		1		1	

Characteristics	**BMI (Kg/m^2)**					
	<25		**25-29.9**		**≥30**	
No	1.17 (0.95-1.44)	0.15	2.96 (2.35-3.72)	0.00	1.90 (1.39-2.60)	0.00
Smoking						
Yes	0.99 (0.64-1.54)	0.96	2.21 (1.39-3.54)	0.00	1.71 (0.88-3.32)	0.12
No	1		1		1	
Physical Activity						
Yes	1		1		1	
No	1.05 (0.81-1.37)	0.69	3.72 (2.68-5.16)	0.00	0.87 (0.59-1.29)	0.49
Menopause status						
Pre-menopause	1		1		1	
Post-menopause	0.85 (0.67-1.06)	0.15	1.27 (1.03-1.55)	0.02	1.72 (1.32-2.25	0.00

There were a total of 426 patients having BMI <25, 543 have BMI 25-29.9 and 269 have BMI ≥30 kg/m^2. The median follow up was 27(95%CI; 27-32), 34(95%CI; 32-34) and 24(95%CI; 24-27) for BMI <25, 25-29.9 and >-30 respectively. We observed no differences in the survival function for BMI categories, $\chi^{12} = 33.3$ with a p-value = 0.0008. Body mass index was clearly not associated with an effect on survival (not necessarily causal) when taken as a continuous variable but it showed significant association when dealt as a categorically variable. We also analyzed if there is a difference in survival functions between BMI categories after adjusting for a potential confounder (Table 4.15 and Figure 4.16).

TABLE 4.15: Survival analysis after adjusting for the covariates.

S. No	Characteristics	HR	P-value	95%CI	Likelihood ratio test	Wald test	Logrank test
BMI as a continuous variable							
1.	BMI	1.01	0.138	0.997-1.032	2.21 on 1df, p=0.137	2.2 on 1 df, p=0.1379	2.2 on 1df, p=0.1378
BMI as a categorical variable							
1.	25-29.9	0.83	0.0211	0.702-0.972	31.56 on 2df, p=0.0001	33.26 on df, p=0.0005	33.88 on 2df, p=0.0004
2.	≥30	1.43	0.0002	1.186-1.731			
Adjusted for covariate age							
1.	25-29.9	0.79	0.0055	0.669-0.933	36.98 on 3df, p=0.0000	38.33 on 3df, p=0.0002	39.02 on 3df, p=0.0000
2.	≥30	1.35	0.0026	1.110-1.639			
3.	Age	0.99	0.0212	0.986-0.999			
Adjusted for covariate Marital status							
1.	25-29.9	0.83	0.0207	0.701- 0.971	31.64 on 3 df, p=0.0006	33.33 on 3df, p=0.0000	33.96 on 3df, p=0.0002
2.	≥30	1.44	0.0002	1.187-1.735			
3.	Un-married	1.02	0.7874	0.882- 1.180			
Adjusted for covariate Number of kids							

S. No	Characteristics	HR	P-value	95%CI	Likelihood ratio test	Wald test	Logrank test
1.	25-29.9	0.83	0.0214	0.702- 0.972	32.97 on 6df, p=0.0001	34.7 on 6df, p=0.0004	35.33 on 6df, p=0.0003
2.	≥30	1.43	0.0003	1.178-1.723			
3.	kids1	1.01	0.9689	0.791- 1.277			
4.	kids2	1.10	0.3878	0.887-1.363			
5.	kids3	0.94	0.6169	0.755- 1.182			
6.	Kids>3	1.05	0.6615	0.856-1.277			
Adjusted for covariate Breast Feeding							
1.	25-29.9	0.83	0.0223	0.703-0.973	31.68 on 3df, p=0.0000	33.37 on 3df, p=0.0000	34 on 3df, p=0.0000
2.	≥30	1.44	0.0002	1.189-1.739			
3.	Breast feeding	0.98	0.7357	0.844-1.127			
Adjusted for covariate Smoking							
1.	25-29.9	0.83	0.0211	0.701-0.971	40.3 on 4df, p=0.0000	41.74 on 4df, p=0.0000	42.6 on 4df, p=0.0001
2.	≥30	1.43	0.0002	1.184-1.727			
3.	Smokers	0.74	0.0141	0.587-0.942			

S. No	Characteristics	HR	P-value	95%CI	Likelihood ratio test	Wald test	Logrank test
Adjusted for covariate Physical Activity							
1.	25-29.9	0.83	0.0202	0.701-0.970	33.54 on 3df, p=0.0000	35.17 on 3df, p=0.0000	35.8 on 3df, p=0.0008
2.	≥30	1.41	0.0004	1.166-1.705			
3.	Physically inactive	1.12	0.1623	0.957-1.304			
Adjusted for covariate Menopause Status							
1.	25-29.9	0.82	0.0188	0.698-0.968	31.92 on 3 df, p=0.0005	33.61 on 3df, p=0.0002	34.23 on 3df, p=0.0000
2.	≥30	1.42	0.0004	1.169-1.717			
3.	Post-menopause	0.96	0.5533	0.829-1.106			

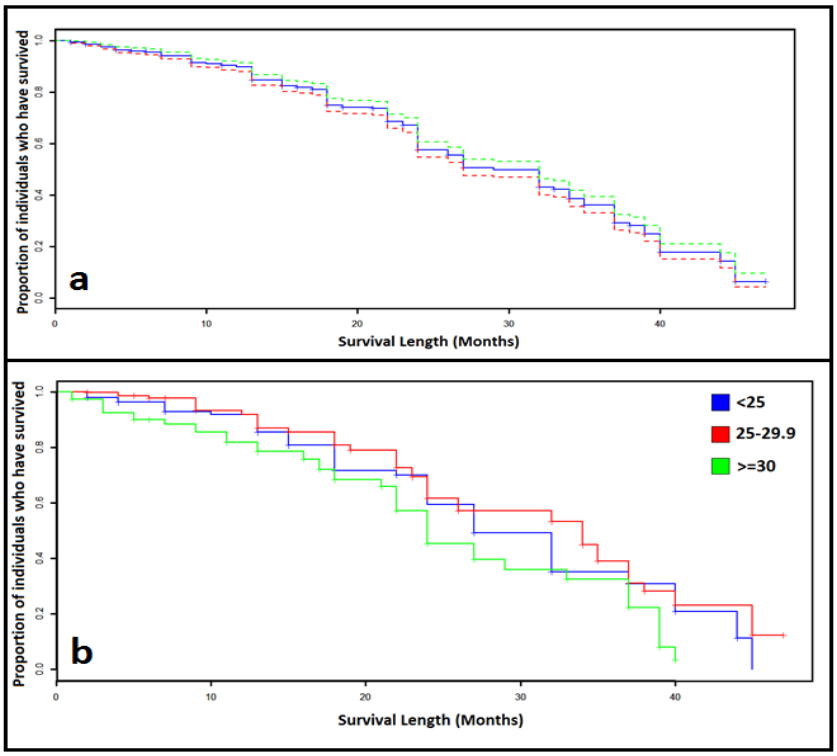

FIGURE 4.16: Plot of Survival curve of female breast cancer patients, (a) Overall Survival curve of female breast cancer patients. Solid blue line represents the survival, while dashed line represents the 95%confidence intervals (b) Survival curve by body mass index.

It is well understood that extra weight gain is associated with an increased incidence of multiple diseases including certain types of cancers. Obesity related morbidity differs among different racial and ethnic groups [235]. This study found that the risk of developing breast cancer increased with increasing BMI at the time of diagnosis in postmenopausal women, highest for BMI ≥30 (OR, 1.72; 95% CI, 1.32-2.25). Our results are consistent with many previous studies [232], [235]–[239] which also states that there is a direct relationship between BMI and risk of breast cancer in postmenopausal women.

Approximately more than 130 epidemiologic studies have [240], [241] studied the association between smoking and breast cancer risk. Despite the presence of a large available data on the topic, there is still no scientific consensus [241], [242], physical inactivity [243] and lower consumption of fruits, vegetables, and fiber [242], all of which are factors known or suspected to increase the risk of cancer

[244], [245]. This study found that the increasing BMI promotes the correlation between smoking and breast cancer risk.

Research showed a relationship between marital status and well being however this connection has not been altogether investigated [207], [209]. Growing literature recommends that psychological elements and the presence or absence of social support might be a vital factor influencing cancer risk [246], [247]. We divided all participating females into married and un-married categories. In terms of marital status, married women were more in control group than cases (75.79% vs. 60.66%) in all BMI categories. Un-married women in all BMI categories are at higher risk than married women. Breast feeding is imagined to decrease the risk of developing breast cancer, which may either due to differentiation of breast tissues or reduction total ovulatory cycles [248]. Another possible system might be, that lactation causes long term endogenous hormonal changes, perhaps reduced estrogen and an increased prolactin secretion, which may decrease a woman's aggregate cumulative exposure to estrogen, subsequently repressing the initiation or growth of breast cancer cells [249]. Non-breast feeding was associated with enhanced breast cancer risk but the risk was not linear with increasing BMI. It showed an increase up to BMI 25-29.9 and then declines in the BMI category $\geq$30Kg/m^2. Children number was found positively associated with breast cancer risk in all BMI categories however, the trend decreases with increase in number of children. It might be because of the summation of all breast feeding episodes or the cumulative length of breast feeding after all births [6].

In this study, we have found an inverse relationship between being overweight (BMI>25.0Kg/m^2) and the overall survival. Similar results for being obese or over-weight at the time of diagnosis has been published in Caucasian populations by a number of other researchers [250]–[253]. To the best of our knowledge, only few studies have explored the correlation between obesity and breast cancer survival in Asians [254], [255], which are not case-control studies. The reason for low survival in overweight patients may be due to the high concentrations of estrogen, testosterone and estradiol. Obesity has additionally been accounted for decreased levels of sex hormone binding globulins which results in an increased level of free

testosterone and estradiol [115]. Therefore, estrogen-sensitive tissues in ladies with BMI more than 25Kg/m^2 are subjected to an increased incitement of estrogen, which may prompt to over development of malignant cells and accordingly advanced metastasis.

In conclusion, the overall trend in association between BMI and malignancy risk is not linear, which might be because of numerous factors including the effect of differential treatment options, and changes in lifestyle after diagnosis. We could not assess this differential effect in the present study due to the limitations of resources. Obesity is associated with poor overall and breast cancer survival. These findings suggest that efforts may be carried out to be in normal BMI range after a breast cancer diagnosis may improve overall survival.

Chapter 5

Conclusion

It was concluded that there is a significant contribution of BRCA1 and BRCA2 genetic alterations and non-genetic risk factors in the pathogenesis of breast cancer among Pakistani women. We believe that this study will be instrumental in setting the genetic components for breast cancer testing, risk assessment and management in this population. The awareness created by this dissertation will also provide a model for providing genetic counseling and other services on etiological basis.

5.1 Summary of Results

We were focusing on the identification of BRCA1 and BRCA2 genetic alterations among Pakistani women affected with Breast Cancer. A total of thirteen variants were detected in BRCA1 and BRCA2 genes respectively. Our reported variants includes three novel variants (Exon3 -37insC, Exon3 -215T<C and Exon14 102-103insTC) in BRCA1 and five novels variants (exon8 +87insA, exon20 +318T<A, exon19 -351-353delTCT, exon16 -17G<T and exon27 T129A) in BRCA2. We also have reported three previously reported SNPs (rs28897686, rs28897696 and rs1060915) in BRCA1 and two in BRCA2 (rs4987049 and rs28897743) along with the novel genetic alterations.

Mean age of cases and controls at recruitment was 50.58±10.68 and 54.78±14.52 years respectively. Cases were comparatively younger than the controls. Mean BMI for the cases and controls were 26.07±4.04 and 25.05±4.25, respectively. Among the patients 60.66% were married, 46.53% were nulliparous, 16.88% had 4 or >4 children, 39.98% females breast fed their children. Considering smoking and physical activity, 88.85% were non-smokers and 67.93% were physically active. Post-menopausal women diagnosed with breast cancer accounted for 52.26%.

The rs28897686 (E1250K) SNP, the novel variants exon3 -37insC and exon3 -215T<C alterations of BRCA1, and rs28897743 (R2336H) SNP, novel variants, exon8 +87insA, exon20 +318T<A, exon19 -351-353delTCT and exon16 -17G<T of BRCA2 were found positively associated the breast cancer risk in the study population. rs28897696 (A1708E) and rs1060915 (S1436S) SNPs of BRCA1 were found non-associated with the disease risk. BRCA1 102-103insTC is exonic variant found in exon14 and was detected only in breast cancer cases. The rs4987049 (Y3308X) SNP introduces a stop codon at position 3308 in BRCA2 and thus results in a truncated protein. T129A mutation is the only reported novel exonic variant of BRCA2 in this study which results in substitution of Aspartic acid to Glutamic acid at position 3260.

There exists a positive association between age and breast cancer. Increase in the basic metabolic index is also correlated with breast cancer and thus the proportion of females having BMI ≥25 was significantly higher among the patients. Overweight (BMI ≥25) and obese (BMI ≥30) females have approximately 1.5 times more risk of having breast cancer. So undoubtedly, obesity is a major risk factor for breast cancer among Pakistani women. Unmarried women were at more than two fold higher risk (OR = 2.03, 95%CI: 1.69-2.44), Nulliparous women had higher risk for breast cancer (OR = 2.56, 95%CI: 1.87-3.51) as compared to parous women. It was further observed that the risk decreases with increase in parity. Use of oral contraceptives (OR = 3.41, 95%CI: 2.86-4.06), and smoking (OR = 1.56, 95%CI: 1.16-2.09) were also significantly associated with increasing risk of breast cancer. Individuals who were physically inactive were recorded to be 1.27 times more likely to develop breast cancer than those who are physically active (OR =

1.75, 95%CI: 1.44-2.12). When the menopausal status was studied, we have found approximately 1.34 fold increase in the disease risk among the postmenopausal patients (OR = 1.34, 95%CI: 1.14-1.58). A total of 269 patients were censored during the study out of the total 1238. The patients had an overall median survival time of 33 months (95%CI: 28-34).

These exploratory analyses indicate that, the studied risk factors were statistically associated with increased risk of breast cancer. It was also observed that mean age at diagnosis is a decade earlier than the western countries.

5.2 Future Directions

On the basis of my personal experience working with breast cancer patients and through a large literature survey following recommendations are made.

1. It is now well understood that the genetic mechanism of breast cancer is highly heterogeneous. There are other genes either dominant or recessive which still needs to be discovered for their role in breast cancer.

2. We do expect more population specific mutations to be revealed.

3. There is still lack of awareness related to breast cancer screening in Pakistan, efforts are needed at the national level to increase awareness among the population especially women. Supportive relationship may be established between the healthcare providers and women for the periodic breast screening which may reduce the likelihood of delayed diagnosis. These efforts will also be helpful at clarifying the misconceptions, fear and embarrassment.

4. Large scale integrative studies are recommended to have better idea about the mechanism of breast cancer development and progression.

Bibliography

[1] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics", *CA: a cancer journal for clinicians*, vol. 61, no. 2, pp. 69–90, 2011.

[2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries", *CA: a cancer journal for clinicians*, 2018.

[3] R. Siegel, E. Ward, O. Brawley, and A. Jemal, "The impact of eliminating socioeconomic and racial disparities on premature cancer deaths", *Ca-a Cancer Journal for Clinicians*, vol. 61, no. 4, pp. 212–236, 2011.

[4] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2015", *CA: a cancer journal for clinicians*, vol. 65, no. 1, pp. 5–29, 2015.

[5] M. D. Althuis, J. M. Dozier, W. F. Anderson, S. S. Devesa, and L. A. Brinton, "Global trends in breast cancer incidence and mortality 1973–1997", *International journal of epidemiology*, vol. 34, no. 2, pp. 405–412, 2005.

[6] M. Singh and B. Jangra, "Association between body mass index and risk of breast cancer among females of north india", *South Asian journal of cancer*, vol. 2, no. 3, p. 121, 2013.

[7] L. A. Torre, F. Islami, R. L. Siegel, E. M. Ward, and A. Jemal, *Global cancer in women: Burden and trends*, 2017.

[8] N. K. Babita, M. Singh, J. S. Malik, and M. Kalhan, "Breastfeeding reduces breast cancer risk: A case-control study in north india", *International journal of preventive medicine*, vol. 5, no. 6, p. 791, 2014.

[9] A. Bener, H. R. El Ayoubi, A. I. Ali, A. Al-Kubaisi, and H. Al-Sulaiti, "Does consanguinity lead to decreased incidence of breast cancer?", *Cancer epidemiology*, vol. 34, no. 4, pp. 413–418, 2010.

[10] P. Hartge, *Genes, hormones, and pathways to breast cancer*, 2003.

[11] B. MacMahon, "Epidemiology and the causes of breast cancer", *International journal of cancer*, vol. 118, no. 10, pp. 2373–2378, 2006.

[12] C. Chung and S. J. Lee, "Estimated risks and optimistic self-perception of breast cancer risk in korean women", *Applied Nursing Research*, vol. 26, no. 4, pp. 180–185, 2013.

[13] F. B. İz and A. Tümer, "Assessment of breast cancer risk and belief in breast cancer screening among the primary healthcare nurses", *Journal of Cancer Education*, vol. 31, no. 3, pp. 575–581, 2016.

[14] S. G. Ayub, S. Rasool, T. Ayub, S. N. Khan, K. A. Wani, and K. I. Andrabi, "Mutational analysis of the brca2 gene in breast carcinoma patients of kashmiri descent", *Molecular medicine reports*, vol. 9, no. 2, pp. 749–753, 2014.

[15] A. L. Calderón-Garcidueñas, P. Ruiz-Flores, R. M. Cerda-Flores, and H. A. Barrera-Saldaña, "Clinical follow up of mexican women with early onset of breast cancer and mutations in the brca1 and brca2 genes", *Salud publica de Mexico*, vol. 47, no. 2, pp. 110–115, 2005.

[16] M.-W. Seong, S. Cho, D.-Y. Noh, W. Han, S.-W. Kim, C.-M. Park, H.-W. Park, S. Kim, J. Kim, and S. Park, "Comprehensive mutational analysis of brca1/brca2 for korean breast cancer patients: Evidence of a founder mutation", *Clinical genetics*, vol. 76, no. 2, pp. 152–160, 2009.

[17] S. E. Martin, M. Sausen, A. Joseph, D. D. Biggs, B. F. Kingham, and E. S. Martin, "Brca1 e1644x: A deleterious mutation in an african american individual with early onset breast cancer", *Breast cancer research and treatment*, vol. 113, no. 2, pp. 393–395, 2009.

[18] E.-H. Lee, S. K. Park, B. Park, S.-W. Kim, M. H. Lee, S. H. Ahn, B. H. Son, K.-Y. Yoo, D. Kang, K. R. Group, *et al.*, "Effect of brca1/2 mutation on short-term and long-term breast cancer survival: A systematic review and meta-analysis", *Breast cancer research and treatment*, vol. 122, no. 1, pp. 11–25, 2010.

[19] C. Heramb, T. Wangensteen, E. M. Grindedal, S. L. Ariansen, S. Lothe, K. R. Heimdal, and L. Mæhle, "Brca1 and brca2 mutation spectrum–an update on mutation distribution in a large cancer genetics clinic in norway", *Hereditary cancer in clinical practice*, vol. 16, no. 1, p. 3, 2018.

[20] S. Chen, E. S. Iversen, T. Friebel, D. Finkelstein, B. L. Weber, A. Eisen, L. E. Peterson, J. M. Schildkraut, C. Isaacs, B. N. Peshkin, *et al.*, "Characterization of brca1 and brca2 mutations in a large united states sample", *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, vol. 24, no. 6, p. 863, Feb. 2006.

[21] D. G. Evans, A. Shenton, E. Woodward, F. Lalloo, A. Howell, and E. R. Maher, "Penetrance estimates for brca1 and brca2 based on genetic testing in a clinical cancer genetics service setting: Risks of breast/ovarian cancer quoted should reflect the cancer burden in the family", *BMC cancer*, vol. 8, no. 1, p. 155, May 2008.

[22] F. P. Liebens, B. Carly, A. Pastijn, and S. Rozenberg, "Management of brca1/2 associated breast cancer: A systematic qualitative review of the state of knowledge in 2006", *European journal of cancer*, vol. 43, no. 2, pp. 238–257, 2007.

[23] A. Antoniou, P. D. Pharoah, S. Narod, H. A. Risch, J. E. Eyfjord, J. L. Hopper, N. Loman, H. Olsson, O. Johannsson, Å. Borg, *et al.*, "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*, vol. 72, no. 5, pp. 1117–1130, 2003.

[24] C. R. James, J. E. Quinn, P. B. Mullan, and P. G. Johnston, *Colin r. james*, Feb. 2007. [Online]. Available: http://theoncologist.alphamedpress.org/content/12/2/142.short.

[25] M. E. Robson, *Treatment of hereditary breast cancer*, Oct. 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0093775407001492.

[26] P. Funchain and A. A. Tarhini, "Using genomic sequencing to improve management in melanoma.", *Oncology (Williston Park, NY)*, vol. 32, no. 3, Mar. 2018.

[27] M. H. Hofker, J. Fu, and C. Wijmenga, *The genome revolution and its role in understanding complex diseases.* Oct. 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24834846.

[28] V. Kuznetsov, H. K. Lee, S. Maurer-Stroh, M. J. Molnár, S. Pongor, B. Eisenhaber, and F. Eisenhaber, "How bioinformatics influences health informatics: Usage of biomolecular sequences, expression profiles and automated microscopic image analyses for clinical needs and public health", *Health Information Science and Systems*, vol. 1, no. 1, p. 2, 2013.

[29] S. A. Narod, *Modifiers of risk of hereditary breast cancer.* Sep. 2006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16998497.

[30] M. A. Clarke and C. E. Joshu, *Early life exposures and adult cancer risk.* Jan. 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/28407101.

[31] J. V. L. J. Lacey, A. R. Kreimer, S. S. Buys, P. M. Marcus, S.-C. Chang, M. F. Leitzmann, R. N. Hoover, P. C. Prorok, C. D. Berg, P. Hartge, and et al., *Breast cancer epidemiology according to recognized breast cancer risk factors in the prostate, lung, colorectal and ovarian (plco) cancer*

screening trial cohort, Mar. 2009. [Online]. Available: https://bmccancer.biomedcentral.com/articles/10.1186/1471-2407-9-84.

[32] V. G. Guerrero, A. F. Baez, C. G. Cofré González, and C. G. Miño González, *Monitoring modifiable risk factors for breast cancer: An obligation for health professionals*, Jun. 2017. [Online]. Available: https://www.scielosp.org/pdf/rpsp/2017.v41/e80/en.

[33] S. Kobayashi, H. Sugiura, Y. Ando, N. Shiraki, T. Yanagi, H. Yamashita, and T. Toyama, *Reproductive history and breast cancer risk*, Jun. 2012. [Online]. Available: https://link.springer.com/article/10.1007/s12282-012-0384-8.

[34] E. F. Beaber, D. S. Buist, W. E. Barlow, K. E. Malone, S. D. Reed, and C. I. Li, Aug. 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4154499/.

[35] K. Bhaskaran, I. Douglas, H. Forbes, I. dos-Santos-Silva, D. A. Leon, and L. Smeeth, *Body-mass index and risk of 22 specific cancers: A population-based cohort study of 5·24 million uk adults*, Aug. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140673614608928.

[36] R. A. Greenberg, B. Sobhian, S. B. Cantor, Y. Nakatani, and D. M. Livingston, *Multifactorial contributions to an acute dna damage response by brca1/bard1-containing complexes*, Jan. 2006. [Online]. Available: http://genesdev.cshlp.org/content/20/1/34.short.

[37] Y. Liu and S. C. West, *Distinct functions of brca1 and brca2 in double-strand break repair*, Dec. 2001. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/11879553.

[38] S. A. Narod and W. D. Foulkes, *Brca1 and brca2: 1994 and beyond*, Sep. 2004. [Online]. Available: https://www.nature.com/articles/nrc1431.

[39] B. Xia, Q. Sheng, K. Nakanishi, A. Ohashi, J. Wu, N. Christ, X. Liu, M. Jasin, F. J. Couch, and D. M. Livingston, "Control of brca2 cellular and clinical functions by a nuclear partner, palb2", *Molecular cell*, vol. 22, no. 6, pp. 719–729, 2006.

[40] G. T. Toh, P. Kang, S. S. Lee, D. S. Lee, S. Y. Lee, S. Selamat, N. A. Mohd, S. Y. Yoon, C. H. Yip, S. H. Teo, and et al., *Brca1 and brca2 germline mutations in malaysian women with early-onset breast cancer without a family history.* Apr. 2008. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18431501.

[41] H. A. Risch, J. R. McLaughlin, D. E. C. Cole, B. Rosen, L. Bradley, E. Kwan, E. Jack, D. J. Vesprini, G. Kuperstein, J. L. A. Abrahamson, and et al., Mar. 2001. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1274482/.

[42] E. Thirthagiri, S. Lee, P. Kang, D. Lee, G. Toh, S. Selamat, S.-Y. Yoon, N. M. Taib, M. Thong, C. Yip, and et al., "Evaluation of brca1 and brca2 mutations and risk-prediction models in a typical asian country (malaysia) with a relatively low incidence of breast cancer", *Breast Cancer Research*, vol. 10, no. 4, R59, Jul. 2008. DOI: 10.1186/bcr2118.

[43] D. R. Youlden, S. M. Cramb, C. H. Yip, and P. D. Baade, "Incidence and mortality of female breast cancer in the asia-pacific region", *Cancer biology & medicine*, vol. 11, no. 2, p. 101, 2014.

[44] M. Zubair, M. T. Khadim, H. Tariq, S. Ali, O. A. Khan, and S. Gul, "Immunohistochemical and clinicopathological factors associated with axillary lymph node metastasis in breast cancer patients of northern pakistan", *Asia Pacific Journal of Cancer Care*, vol. 2, no. 4, 2018.

[45] K. Usmani, A. Khanum, H. Afzal, and N. Ahmad, "Breast carcinoma in pakistani women", *Journal of environmental pathology, toxicology and oncology : official organ of the International Society for Environmental Toxicology and Cancer*, vol. 15, no. 2-4, pp. 251–253, 1996, ISSN: 0731-8898. [Online]. Available: http://europepmc.org/abstract/MED/9216816.

[46] E. CHYip, "Breast cancer-a comparative study between malaysian and singaporean women", *Singapore Med J*, vol. 37, pp. 264–267, 1996.

[47] H. Iraj, K. Mojgan, K. Amir, and J. M. Amir, "Breast cancer in iran: Result of a multi- center study", vol. 5, pp. 24–27, 2004.

[48] A. N. Hisham and C.-H. Yip, "Overview of breast cancer in malaysian women: A problem with late diagnosis", *Asian Journal of Surgery*, vol. 27, no. 2, pp. 130–133, 2004, ISSN: 1015-9584. DOI: https://doi.org/10.1016/S1015-9584(09)60326-2. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1015958409603262.

[49] A. Nissan, R. M. Spira, T. Hamburger, M. Badrriyah, D. Prus, T. Cohen, A. Hubert, H. R. Freund, and T. Peretz, "Clinical profile of breast cancer in arab and jewish women in the jerusalem area", *The American journal of surgery*, vol. 188, no. 1, pp. 62–67, 2004.

[50] R. Menhas and S. Umer, "Breast cancer among pakistani women.", *Iranian journal of public health*, vol. 44, no. 4, pp. 586–587, 2015.

[51] H. M. Asif, S. Sultana, N. Akhtar, J. U. Rehman, and R. U. Rehman, "Prevalence, risk factors and disease knowledge of breast cancer in pakistan", *Asian Pac J Cancer Prev*, vol. 15, no. 11, pp. 4411–6, 2014.

[52] R. Pervaiz, "Genetic mutations associated with breast cancer in pakistan", *Malaysian Journal of Medical and Biological Research*, vol. 2, no. 3, pp. 308–313, 2015.

[53] Y. Bhurgri, "Karachi cancer registry data–implications for the national cancer control program of pakistan", *Asian Pac J Cancer Prev*, vol. 5, no. 1, pp. 77–82, 2004.

[54] S. P. Leong, Z.-Z. Shen, T.-J. Liu, G. Agarwal, T. Tajima, N.-S. Paik, K. Sandelin, A. Derossis, H. Cody, and W. D. Foulkes, "Is breast cancer the same disease in asian and western countries?", *World journal of surgery*, vol. 34, no. 10, pp. 2308–2324, 2010.

[55] K. N. Lai, W. K. Ho, I. N. Kang, P. C. E. Kang, S. Y. Phuah, S. Mariapun, C.-H. Yip, N. A. M. Taib, and S.-H. Teo, "Characterization of brca1 and brca2 variants in multi-ethnic asian cohort from a malaysian case-control study", *BMC cancer*, vol. 17, no. 1, p. 149, 2017.

[56] J. E. Joy, E. E. Penhoet, and D. B. Petitti, *Saving women's lives: Strategies for improving breast cancer detection and diagnosis*. 2005.

[57] M. G. Valsecchi and E. Steliarova-Foucher, "Cancer registration in developing countries: Luxury or necessity?", *The lancet oncology*, vol. 9, no. 2, pp. 159–167, 2008.

[58] T. R. Rebbeck, T. M. Friebel, E. Friedman, U. Hamann, D. Huo, A. Kwong, E. Olah, O. I. Olopade, A. R. Solano, S.-H. Teo, *et al.*, "Mutational spectrum in a worldwide study of 29,700 families with brca1 or brca2 mutations", *Human mutation*, vol. 39, no. 5, pp. 593–620, 2018.

[59] J. Pérez-Mayoral, A. Cora-Morges, K. Gonzalez-Rosa, J. Hernandez, D. Castillo, J. Herzog, J. Weitzel, and M. Cruz-Correa, *Abstract b53: Brca1 and brca2 mutations spectrum in puerto rican hispanics*, 2017.

[60] I. Sakamoto, Y. Hirotsu, H. Nakagomi, H. Ouchi, A. Ikegami, K. Teramoto, K. Amemiya, H. Mochizuki, and M. Omata, "Brca 1 and brca 2 mutations in j apanese patients with ovarian, fallopian tube, and primary peritoneal cancer", *Cancer*, vol. 122, no. 1, pp. 84–90, 2016.

[61] V. Silvestri, D. Barrowdale, A. M. Mulligan, S. L. Neuhausen, S. Fox, B. Y. Karlan, G. Mitchell, P. James, D. L. Thull, K. K. Zorn, *et al.*, "Male breast cancer in brca1 and brca2 mutation carriers: Pathology data from the consortium of investigators of modifiers of brca1/2", *Breast Cancer Research*, vol. 18, no. 1, p. 15, 2016.

[62] E. Walsh, M. Farrell, C. Nolan, F. Gallagher, R. Clarke, J. McCaffrey, M. Kennedy, M. Barry, M. Kell, and D. Gallagher, "Breast cancer detection among irish brca1 & brca2 mutation carriers: A population-based study", *Irish Journal of Medical Science (1971-)*, vol. 185, no. 1, pp. 189–194, 2016.

[63] F. Quiles, À. Teulé, N. Martinussen Tandstad, L. Feliubadaló, E. Tornero, J. Del Valle, M. Menéndez, M. Salinas, V. Wethe Rognlien, A. Velasco, *et al.*, "Identification of a founder brca1 mutation in the moroccan population", *Clinical genetics*, vol. 90, no. 4, pp. 361–365, 2016.

[64] N. Tung, C. Battelli, B. Allen, R. Kaldate, S. Bhatnagar, K. Bowles, K. Timms, J. E. Garber, C. Herold, L. Ellisen, *et al.*, "Frequency of mutations in individuals with breast cancer referred for brca 1 and brca 2 testing using

next-generation sequencing with a 25-gene panel", *Cancer*, vol. 121, no. 1, pp. 25–33, 2015.

[65] T. R. Rebbeck, N. Mitra, F. Wan, O. M. Sinilnikova, S. Healey, L. McGuffog, S. Mazoyer, G. Chenevix-Trench, D. F. Easton, A. C. Antoniou, *et al.*, "Association of type and location of brca1 and brca2 mutations with risk of breast and ovarian cancer", *JAMA*, vol. 313, no. 13, pp. 1347–1361, 2015.

[66] C. Villarreal-Garza, J. Weitzel, M. Llacuachaqui, E. Sifuentes, M. Magallanes-Hoyos, L. Gallardo, R. Alvarez-Gómez, J. Herzog, D. Castillo, R. Royer, *et al.*, "The prevalence of brca1 and brca2 mutations among young mexican women with triple-negative breast cancer", *Breast cancer research and treatment*, vol. 150, no. 2, pp. 389–394, 2015.

[67] J. Mersch, M. A. Jackson, M. Park, D. Nebgen, S. K. Peterson, C. Singletary, B. K. Arun, and J. K. Litton, "Cancers associated with brca 1 and brca 2 mutations other than breast and ovarian", *Cancer*, vol. 121, no. 2, pp. 269–275, 2015.

[68] B. Norquist, M. Harrell, T. Walsh, J. Mandell, K. Agnew, M. Lee, K. Pennington, M. King, and E. Swisher, "Germline mutations in cancer susceptibility genes in brca1 and brca2 negative families with ovarian and breast cancer", *Gynecologic Oncology*, vol. 135, no. 2, p. 383, 2014.

[69] M. Akbari, T. Donenberg, J. Lunn, D. Curling, T. Turnquest, E. Krill-Jackson, S. Zhang, S. Narod, and J. Hurley, "The spectrum of brca1 and brca2 mutations in breast cancer patients in the bahamas", *Clinical genetics*, vol. 85, no. 1, pp. 64–67, 2014.

[70] S. A. Narod and L. Salmena, "Brca1 and brca2 mutations and breast cancer", *Discovery medicine*, vol. 12, no. 66, pp. 445–453, Nov. 2011.

[71] Y. Laitman, R. T. Borsthein, D. Stoppa-Lyonnet, E. Dagan, L. Castera, M. Goislard, R. Gershoni-Baruch, H. Goldberg, B. Kaufman, N. Ben-Baruch, *et al.*, "Germline mutations in brca1 and brca2 genes in ethnically diverse high risk families in israel", *Breast cancer research and treatment*, vol. 127, no. 2, pp. 489–495, Jun. 2011.

[72] L. Cavallone, S. L. Arcand, C. M. Maugard, S. Nolet, L. A. Gaboury, A.-M. Mes-Masson, P. Ghadirian, D. Provencher, and P. N. Tonin, "Comprehensive brca1 and brca2 mutation analyses and review of french canadian families with at least three cases of breast cancer", *Familial cancer*, vol. 9, no. 4, pp. 507–517, Dec. 2010.

[73] J. D. Fackenthal and O. I. Olopade, "Breast cancer risk associated with brca1 and brca2 in diverse populations", *Nature Reviews Cancer*, vol. 7, no. 12, p. 937, Dec. 2007.

[74] J. Hartikainen, V. Kataja, M. Pirskanen, A. Arffman, U. Ristonmaa, P. Vahteristo, M. Ryynänen, S. Heinonen, V.-M. Kosma, and A. Mannermaa, "Screening for brca1 and brca2 mutations in eastern finnish breast/ovarian cancer families", *Clinical genetics*, vol. 72, no. 4, pp. 311–320, Oct. 2007.

[75] S. H. Ahn, B. H. Son, K.-S. Yoon, D.-Y. Noh, W. Han, S.-W. Kim, E. S. Lee, H.-L. Park, Y. J. Hong, J. J. Choi, *et al.*, "Brca1 and brca2 germline mutations in korean breast cancer patients at high risk of carrying mutations", *Cancer letters*, vol. 245, no. 1-2, pp. 90–95, Jan. 2007.

[76] A. Woodward, T. Davis, A. Silva, J. Kirk, and J. Leary, "Large genomic rearrangements of both brca2 and brca1 are a feature of the inherited breast-/ovarian cancer phenotype in selected families", *Journal of medical genetics*, vol. 42, no. 5, e31–e31, May 2005.

[77] A. Nadine, D. Easton, J. Chang-Claude, M. Rookus, R. Brohet, E. Cardis, A. Antoniou, T. Wagner, J. Simard, G. Evans, *et al.*, "Effect of chest x-rays on the risk of breast cancer among brca1/2 mutation carriers in the international brca1/2 carrier cohort study: A report from the embrace, genepso, geo-hebon, and ibccs collaborators' group.", *J Clin Oncol*, vol. 24, no. 21, pp. 3361–6, Jul. 2006.

[78] P. Pohlreich, M. Zikan, J. Stribrna, Z. Kleibl, M. Janatova, J. Kotlas, J. Zidovska, J. Novotny, L. Petruzelka, C. Szabo, *et al.*, "High proportion of recurrent germline mutations in the brca1 gene in breast and ovarian cancer

patients from the prague area", *Breast cancer research*, vol. 7, no. 5, R728, Jul. 2005.

[79] K. Claes, B. Poppe, E. Machackova, I. Coene, L. Foretova, A. De Paepe, and L. Messiaen, "Differentiating pathogenic mutations from polymorphic alterations in the splice sites of brca1 and brca2", *Genes, Chromosomes and Cancer*, vol. 37, no. 3, pp. 314–320, Apr. 2003.

[80] A. Liede, I. A. Malik, Z. Aziz, P. De los Rios, E. Kwan, and S. A. Narod, "Contribution of brca1 and brca2 mutations to breast and ovarian cancer in pakistan", *The American Journal of Human Genetics*, vol. 71, no. 3, pp. 595–606, Sep. 2002.

[81] T. Peelen, M. Van Vliet, A. Petrij-Bosch, R. Mieremet, C. Szabo, A. Van den Ouweland, F. Hogervorst, R. Brohet, M. Ligtenberg, E. Teugels, *et al.*, "A high proportion of novel mutations in brca1 with strong founder effects among dutch and belgian hereditary breast and ovarian cancer families", *American journal of human genetics*, vol. 60, no. 5, p. 1041, May 1997.

[82] P. Vehmanen, L. S. Friedman, H. Eerola, M. McClure, B. Ward, L. Sarantaus, T. Kainu, K. Syrjäkoski, S. Pyrhönen, O.-P. Kallioniemi, *et al.*, "Low proportion of brca1 and brca2 mutations in finnish breast cancer families: Evidence for additional susceptibility genes", *Human molecular genetics*, vol. 6, no. 13, pp. 2309–2315, Dec. 1997.

[83] G.-P. Xu, Q. Zhao, D. Wang, W.-Y. Xie, L.-J. Zhang, H. Zhou, S.-Z. Chen, and L.-F. Wu, "The association between brca1 gene polymorphism and cancer risk: A meta-analysis", *Oncotarget*, vol. 9, no. 9, p. 8681, 2018.

[84] Y. Zhu, J. K. Kan Zhai, J. Li, Y. Gong, Y. Yang, J. Tian, Y. Zhang, D. Zou, X. Peng, J. Gong, *et al.*, "Brca1 missense polymorphisms are associated with poor prognosis of pancreatic cancer patients in a chinese population", *Oncotarget*, vol. 8, no. 22, p. 36 033, 2017.

[85] M. C. Prosperi, S. L. Ingham, A. Howell, F. Lalloo, I. E. Buchan, and D. G. Evans, "Can multiple snp testing in brca2 and brca1 female carriers be used to improve risk prediction models in conjunction with clinical assessment?",

BMC medical informatics and decision making, vol. 14, no. 1, p. 87, Oct. 2014.

[86] A. Juwle and D. Saranath, "Brca1/brca2 gene mutations/snps and brca1 haplotypes in early-onset breast cancer patients of indian ethnicity", *Medical oncology*, vol. 29, no. 5, pp. 3272–3281, Dec. 2012.

[87] R. Milne and A. Antoniou, "Genetic modifiers of cancer risk for brca1 and brca2 mutation carriers", *Annals of oncology*, vol. 22, no. suppl_1, pp. i11–i17, Jan. 2011.

[88] M. Pongsavee, V. Yamkamon, S. Dakeng, P. O-charoenrat, D. R. Smith, G. F. Saunders, and P. Patmasiriwat, "The brca1 3-utr: 5711+ 421t/t_5711+ 1286t/t genotype is a possible breast and ovarian cancer risk factor", *Genetic testing and molecular biomarkers*, vol. 13, no. 3, pp. 307–317, 2009.

[89] C. Pelletier, W. C. Speed, T. Paranjape, K. Keane, R. Blitzblau, A. Hollestelle, K. Safavi, A. van den Ouweland, D. Zelterman, F. J. Slack, *et al.*, "Rare brca1 haplotypes including 3'utr snps associated with breast cancer risk", *Cell Cycle*, vol. 10, no. 1, pp. 90–99, Jan. 2011.

[90] A. C. Antoniou, J. Beesley, L. McGuffog, O. M. Sinilnikova, S. Healey, S. L. Neuhausen, Y. C. Ding, T. R. Rebbeck, J. N. Weitzel, H. T. Lynch, and et al., *Common breast cancer susceptibility alleles and the risk of breast cancer for brca1 and brca2 mutation carriers: Implications for risk prediction*, Dec. 2010.

[91] Antoniou, A. C., O. M., McGuffog, Lesley, Healey, Heli, Heikkinen, Simard, Jacques, and et al., *Common variants in lsp1 , 2q35 and 8q24 and breast cancer risk for brca1 and brca2 mutation carriers*, Aug. 2009. [Online]. Available: https://academic.oup.com/hmg/article/18/22/4442/587713.

[92] A. C. Antoniou, A. B. Spurdle, O. M. Sinilnikova, S. Healey, K. A. Pooley, R. K. Schmutzler, B. Versmold, C. Engel, A. Meindl, N. Arnold, *et al.*, "Common breast cancer-predisposition alleles are associated with breast

cancer risk in brca1 and brca2 mutation carriers", *The American Journal of Human Genetics*, vol. 82, no. 4, pp. 937–948, Apr. 2008.

[93] A. C. Antoniou, O. M. Sinilnikova, J. Simard, M. Léoné, M. Dumont, S. L. Neuhausen, J. P. Struewing, D. Stoppa-Lyonnet, L. Barjhoux, D. J. Hughes, *et al.*, "Rad51 135g c modifies breast cancer risk among brca2 mutation carriers: Results from a combined analysis of 19 studies", *The American Journal of Human Genetics*, vol. 81, no. 6, pp. 1186–1200, Dec. 2007.

[94] M. Antony, B. Surakutty, T. Vasu, and M. Chisthi, "Risk factors for breast cancer among indian women: A case–control study", *Nigerian journal of clinical practice*, vol. 21, no. 4, Apr. 2018.

[95] B. Rosner, A. H. Eliassen, A. T. Toriola, W. Y. Chen, S. E. Hankinson, W. C. Willett, C. S. Berkey, and G. A. Colditz, "Weight and weight changes in early adulthood and later breast cancer risk", *International journal of cancer*, vol. 140, no. 9, pp. 2003–2014, May 2017.

[96] S. H. Nelson, C. R. Marinac, R. E. Patterson, S. J. Nechuta, S. W. Flatt, B. J. Caan, M. L. Kwan, E. M. Poole, W. Y. Chen, X.-o. Shu, *et al.*, "Impact of very low physical activity, bmi, and comorbidities on mortality among breast cancer survivors", *Breast cancer research and treatment*, vol. 155, no. 3, pp. 551–557, Feb. 2016.

[97] M. N. Passarelli, P. A. Newcomb, J. M. Hampton, A. Trentham-Dietz, L. J. Titus, K. M. Egan, J. A. Baron, and W. C. Willett, "Cigarette smoking before and after breast cancer diagnosis: Mortality from breast cancer and smoking-related diseases", *Journal of Clinical Oncology*, vol. 34, no. 12, p. 1315, Apr. 2016.

[98] M. M. Gaudet, S. M. Gapstur, J. Sun, W. R. Diver, L. M. Hannan, and M. J. Thun, "Active smoking and breast cancer risk: Original cohort data and meta-analysis", *Journal of the National Cancer Institute*, vol. 105, no. 8, pp. 515–525, Apr. 2013.

[99] D. Chan, A. Vieira, D. Aune, E. Bandera, D. Greenwood, A. McTiernan, D. Navarro Rosenblatt, I. Thune, R. Vieira, and T. Norat, "Body mass index and survival in women with breast cancer—systematic literature review and meta-analysis of 82 follow-up studies", *Annals of Oncology*, vol. 25, no. 10, pp. 1901–1914, Oct. 2014.

[100] S. Berube, J. Lemieux, L. Moore, E. Maunsell, and J. Brisson, "Smoking at time of diagnosis and breast cancer-specific survival: New findings and systematic review with meta-analysis", *Breast Cancer Research*, vol. 16, no. 2, R42, Apr. 2014.

[101] L. Dossus, M.-C. Boutron-Ruault, R. Kaaks, I. T. Gram, A. Vilier, B. Fervers, J. Manjer, A. Tjonneland, A. Olsen, K. Overvad, *et al.*, "Active and passive cigarette smoking and breast cancer risk: Results from the epic cohort", *International journal of cancer*, vol. 134, no. 8, pp. 1871–1888, Apr. 2014.

[102] A. Young, E. Weltzien, M. Kwan, A. Castillo, B. Caan, and C. H. Kroenke, "Pre-to post-diagnosis weight change and associations with physical functional limitations in breast cancer survivors", *Journal of Cancer Survivorship*, vol. 8, no. 4, pp. 539–547, 2014.

[103] Z. J. Andersen, J. L. Baker, K. Bihrmann, I. Vejborg, T. I. Sorensen, and E. Lynge, "Birth weight, childhood body mass index, and height in relation to mammographic density and breast cancer: A register-based cohort study", *Breast Cancer Research*, vol. 16, no. 1, R4, Jan. 2014.

[104] "Effect of obesity in premenopausal er+ early breast cancer: Ebctcg data on 80,000 patients in 70 trials", *Journal of Clinical Oncology*, vol. 32, no. 15_suppl, pp. 503–503, May 2014.

[105] R. Ritte, A. Lukanova, A. Tjønneland, A. Olsen, K. Overvad, S. Mesrine, G. Fagherazzi, L. Dossus, B. Teucher, K. Steindorf, *et al.*, "Height, age at menarche and risk of hormone receptor-positive and-negative breast cancer: A cohort study", *International journal of cancer*, vol. 132, no. 11, pp. 2619–2629, Jun. 2013.

[106] M. D. Alsaker, S. Opdahl, P. R. Romundstad, and L. J. Vatten, "Association of time since last birth, age at first birth and parity with breast cancer survival among parous women: A register-based study from norway", *International journal of cancer*, vol. 132, no. 1, pp. 174–181, Jan. 2013.

[107] N. Biglia, E. Peano, P. Sgandurra, G. Moggio, S. Pecchio, F. Maggiorotto, and P. Sismondi, "Body mass index (bmi) and breast cancer: Impact on tumor histopatologic features, cancer subtypes and recurrence rate in pre and postmenopausal women", *Gynecological Endocrinology*, vol. 29, no. 3, pp. 263–267, Mar. 2013.

[108] J. Horn, B. O. Asvold, S. Opdahl, S. Tretli, and L. J. Vatten, "Reproductive factors and the risk of breast cancer in old age: A norwegian cohort study", *Breast cancer research and treatment*, vol. 139, no. 1, pp. 237–243, May 2013.

[109] R. Scheri, S. Power, J. Marks, V. Seewaldt, K. Marcom, and S. Hwang, "Abstract p4-13-06: Association of age, obesity and incident breast cancer phenotypes", *Cancer Research*, vol. 72, no. 24 Supplement, P4-13-06, Dec. 2012.

[110] W. van de Water, C. Markopoulos, C. J. van de Velde, C. Seynaeve, A. Hasenburg, D. Rea, H. Putter, J. W. Nortier, A. J. de Craen, E. T. Hille, *et al.*, "Association between age at diagnosis and disease-specific mortality among postmenopausal women with hormone receptor–positive breast cancer", *JAMA*, vol. 307, no. 6, pp. 590–597, Feb. 2012.

[111] M. Kawai, Y. Minami, Y. Nishino, K. Fukamachi, N. Ohuchi, and Y. Kakugawa, "Body mass index and survival after breast cancer diagnosis in japanese women", *BMC cancer*, vol. 12, no. 1, p. 149, Apr. 2012.

[112] C. G. on Hormonal Factors in Breast Cancer *et al.*, "Menarche, menopause, and breast cancer risk: Individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies", *The lancet oncology*, vol. 13, no. 11, pp. 1141–1151, Nov. 2012.

[113] S. Paluch-Shimon, I. Wolf, S. Sadetzki, I. Gluck, B. Oberman, M. Z. Papa, R. Catane, and B. Kaufman, "Association between very young age and adverse characteristics of breast cancer at presentation amongst israeli women", *American journal of clinical oncology*, vol. 34, no. 3, pp. 219–222, Jun. 2011.

[114] S. E. Singletary, "Rating the risk factors for breast cancer", *Annals of surgery*, vol. 237, no. 4, p. 474, 2003.

[115] A. McTiernan, "Behavioral risk factors in breast cancer: Can risk be modified?", *The oncologist*, vol. 8, no. 4, pp. 326–334, Aug. 2003.

[116] J. Sambrook and D. W. Russell, "Purification of nucleic acids by extraction with phenol: Chloroform", *Cold Spring Harbor Protocols*, vol. 2006, no. 1, pdb–prot4455, Jun. 2006.

[117] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen, "Primer3—new capabilities and interfaces", *Nucleic acids research*, vol. 40, no. 15, e115–e115, Aug. 2012.

[118] H. Yan, Z. Dobbie, S. B. Gruber, S. Markowitz, K. Romans, F. M. Giardiello, K. W. Kinzler, and B. Vogelstein, "Small changes in expression affect predisposition to tumorigenesis", *Nature genetics*, vol. 30, no. 1, p. 25, Jan. 2002.

[119] S. Benhamou and A. Sarasin, "Ercc2/xpd gene polymorphisms and cancer risk", *Mutagenesis*, vol. 17, no. 6, pp. 463–469, Nov. 2002.

[120] F. Durocher, D. Shattuck-Eidens, M. McClure, F. Labrie, M. H. Skolnick, D. E. Goldgar, and J. Simard, "Comparison of brca1 polymorphisms, rare sequence variants and/or missense mutations in unaffected and breast/ovarian cancer populations", *Human molecular genetics*, vol. 5, no. 6, pp. 835–842, Jun. 1996.

[121] L.-X. Qiu, L. Yao, K. Xue, J. Zhang, C. Mao, B. Chen, P. Zhan, H. Yuan, and X.-C. Hu, "Brca2 n372h polymorphism and breast cancer susceptibility: A meta-analysis involving 44,903 subjects", *Breast cancer research and treatment*, vol. 123, no. 2, pp. 487–490, Sep. 2010.

[122] W.-Q. Xue, Y.-Q. He, J.-H. Zhu, J.-Q. Ma, J. He, and W.-H. Jia, "Association of brca2 n372h polymorphism with cancer susceptibility: A comprehensive review and meta-analysis", *Scientific reports*, vol. 4, p. 6791, 2014.

[123] Q. Li, R. Guan, Y. Qiao, C. Liu, N. He, X. Zhang, X. Jia, H. Sun, J. Yu, and L. Xu, "Association between the brca2 rs144848 polymorphism and cancer susceptibility: A meta-analysis", *Oncotarget*, vol. 8, no. 24, p. 39 818, Jun. 2017.

[124] C. S. Healey, A. M. Dunning, M. D. Teare, D. Chase, L. Parker, J. Burn, J. Chang-Claude, A. Mannermaa, V. Kataja, D. G. Huntsman, *et al.*, "A common variant in brca2 is associated with both breast cancer risk and prenatal viability", *Nature genetics*, vol. 26, no. 3, p. 362, Nov. 2000.

[125] A. B. Spurdle, J. L. Hopper, X. Chen, G. S. Dite, J. Cui, M. R. McCredie, G. G. Giles, S. Ellis-Steinborner, D. J. Venter, B. Newman, *et al.*, "The brca2 372 hh genotype is associated with risk of breast cancer in australian women under age 60 years", *Cancer Epidemiology and Prevention Biomarkers*, vol. 11, no. 4, pp. 413–416, Apr. 2002.

[126] D. A. Hill, S. S. Wang, J. R. Cerhan, S. Davis, W. Cozen, R. K. Severson, P. Hartge, S. Wacholder, M. Yeager, S. J. Chanock, *et al.*, "Risk of non-hodgkin lymphoma (nhl) in relation to germline variation in dna repair and related genes", *Blood*, vol. 108, no. 9, pp. 3161–3167, Nov. 2006.

[127] S. T. Sherry, M.-H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, and K. Sirotkin, "Dbsnp: The ncbi database of genetic variation", *Nucleic acids research*, vol. 29, no. 1, pp. 308–311, 2001.

[128] R. Karchin, "Next generation tools for the annotation of human snps", *Briefings in bioinformatics*, vol. 10, no. 1, pp. 35–52, Jan. 2009.

[129] P. C. Ng and S. Henikoff, "Predicting the effects of amino acid substitutions on protein function", *Annu. Rev. Genomics Hum. Genet.*, vol. 7, pp. 61–80, Sep. 2006.

[130] S. Mooney, "Bioinformatics approaches and resources for single nucleotide polymorphism functional analysis", *Briefings in bioinformatics*, vol. 6, no. 1, pp. 44–56, Mar. 2005.

[131] R. E. Steward, M. W. MacArthur, R. A. Laskowski, and J. M. Thornton, "Molecular basis of inherited diseases: A structural perspective", *TRENDS in Genetics*, vol. 19, no. 9, pp. 505–513, Sep. 2003.

[132] P. D. Thomas, M. J. Campbell, A. Kejariwal, H. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan, and A. Narechania, "Panther: A library of protein families and subfamilies indexed by function", *Genome research*, vol. 13, no. 9, pp. 2129–2141, Sep. 2003.

[133] N.-L. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider, and P. C. Ng, "Sift web server: Predicting effects of amino acid substitutions on proteins", *Nucleic acids research*, vol. 40, no. W1, W452–W457, 2012.

[134] H. Mi, X. Huang, A. Muruganujan, H. Tang, C. Mills, D. Kang, and P. D. Thomas, "Panther version 11: Expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements", *Nucleic acids research*, vol. 45, no. D1, pp. D183–D189, Jan. 2016.

[135] B. Reva, Y. Antipin, and C. Sander, "Predicting the functional impact of protein mutations: Application to cancer genomics", *Nucleic acids research*, vol. 39, no. 17, e118–e118, Sep. 2011.

[136] Y. Choi and A. P. Chan, "Provean web server: A tool to predict the functional effect of amino acid substitutions and indels", *Bioinformatics*, vol. 31, no. 16, pp. 2745–2747, Aug. 2015.

[137] M. De Jong, I. Nolte, G. Te Meerman, W. Van der Graaf, J. Oosterwijk, J. Kleibeuker, M. Schaapveld, and E. De Vries, "Genes other than brca1 and brca2 involved in breast cancer susceptibility", *Journal of medical genetics*, vol. 39, no. 4, pp. 225–242, Apr. 2002.

[138] W. Burke, M. Daly, J. Garber, J. Botkin, M. J. E. Kahn, P. Lynch, A. McTiernan, K. Offit, J. Perlman, G. Petersen, *et al.*, "Recommendations for follow-up care of individuals with an inherited predisposition to cancer: Ii. brca1 and brca2", *Jama*, vol. 277, no. 12, pp. 997–1003, Mar. 1997.

[139] E. L. Schubert, M. K. Lee, H. C. Mefford, R. H. Argonza, J. E. Morrow, J. Hull, J. L. Dann, and M.-C. King, "Brca2 in american families with four or more cases of breast or ovarian cancer: Recurrent and novel mutations, variable expression, penetrance, and the possibility of families whose cancer is not attributable to brca1 or brca2.", *American journal of human genetics*, vol. 60, no. 5, p. 1031, May 1997.

[140] X. Solé, E. Guinó, J. Valls, R. Iniesta, and V. Moreno, "Snpstats: A web tool for the analysis of association studies", *Bioinformatics*, vol. 22, no. 15, pp. 1928–1929, 2006.

[141] J. Shen, Z. Li, J. Chen, Z. Song, Z. Zhou, and Y. Shi, "Shesisplus, a toolset for genetic studies on polyploid species", *Scientific reports*, vol. 6, p. 24 095, 2016.

[142] Y. Yong and H. Lin, "Shesis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci", *Cell research*, vol. 15, no. 2, p. 97, 2005.

[143] L. K. Mahan, S. Escott-Stump, *et al.*, *Krause's food, nutrition, & diet therapy*. Saunders Philadelphia, 2004, vol. 11.

[144] R. C. Team *et al.*, "R: A language and environment for statistical computing", 2014.

[145] J. T. Rich, J. G. Neely, R. C. Paniello, C. C. Voelker, B. Nussenbaum, and E. W. Wang, "A practical guide to understanding kaplan-meier curves", *Otolaryngology—Head and Neck Surgery*, vol. 143, no. 3, pp. 331–336, 2010.

[146] O. Dıez, A. Osorio, M. Durán, J. I. Martinez-Ferrandis, M. d. l. Hoya, R. Salazar, A. Vega, B. Campos, R. Rodrıguez-López, E. Velasco, *et al.*,

"Analysis of brca1 and brca2 genes in spanish breast/ovarian cancer patients: A high proportion of mutations unique to spain and evidence of founder effects", *Human mutation*, vol. 22, no. 4, pp. 301–312, Oct. 2003.

[147] L. Jara, S. Ampuero, E. Santibañez, L. Seccia, J. Rodrıguez, M. Bustamante, G. Lay-Son, J. OJEDA, J. M. Reyes, and R. Blanco, "Molecular analysis of the eighteen most frequent mutations in the brca1 gene in 63 chilean breast cancer families", *Biological research*, vol. 37, no. 3, pp. 469–481, 2004.

[148] P. E. Bonnen, P. J. Wang, M. Kimmel, R. Chakraborty, and D. L. Nelson, "Haplotype and linkage disequilibrium architecture for human cancer-associated genes", *Genome research*, vol. 12, no. 12, pp. 1846–1853, Dec. 2002.

[149] A. M. Dunning, M. Chiano, N. R. Smith, J. Dearden, M. Gore, S. Oakes, C. Wilson, M. Stratton, J. Peto, D. Easton, *et al.*, "Common brca1 variants and susceptibility to breast and ovarian cancer in the general population", *Human molecular genetics*, vol. 6, no. 2, pp. 285–289, Feb. 1997.

[150] D. G. Cox, P. Kraft, S. E. Hankinson, and D. J. Hunter, "Haplotype analysis of common variants in the brca1 gene and risk of sporadic breast cancer", *Breast Cancer Research*, vol. 7, no. 2, R171, Dec. 2005.

[151] M. L. Freedman, K. L. Penney, D. O. Stram, L. Le Marchand, J. N. Hirschhorn, L. N. Kolonel, D. Altshuler, B. E. Henderson, and C. A. Haiman, "Common variation in brca2 and breast cancer risk: A haplotype-based analysis in the multiethnic cohort", *Human molecular genetics*, vol. 13, no. 20, pp. 2431–2441, Oct. 2004.

[152] C. Baynes, C. S. Healey, K. A. Pooley, S. Scollen, R. N. Luben, D. J. Thompson, P. D. Pharoah, D. F. Easton, B. A. Ponder, and A. M. Dunning, "Common variants in the atm, brca1, brca2, chek2 and tp53 cancer susceptibility genes are unlikely to increase breast cancer risk", *Breast cancer research*, vol. 9, no. 2, R27, Apr. 2007.

[153] K. Claes, B. Poppe, I. Coene, A. De Paepe, and L. Messiaen, "Brca1 and brca2 germline mutation spectrum and frequencies in belgian breast/ovarian cancer families", *British journal of cancer*, vol. 90, no. 6, p. 1244, Mar. 2004.

[154] S. Hussain, J. B. Wilson, A. L. Medhurst, J. Hejna, E. Witt, S. Ananth, A. Davies, J.-Y. Masson, R. Moses, S. C. West, *et al.*, "Direct interaction of fancd2 with brca2 in dna damage response pathways", *Human molecular genetics*, vol. 13, no. 12, pp. 1241–1248, Jun. 2004.

[155] M. Janatova, M. Zikan, P. Dundr, B. Matous, and P. Pohlreich, "Novel somatic mutations in the brca1 gene in sporadic breast tumors", *Human mutation*, vol. 25, no. 3, pp. 319–319, Feb. 2005.

[156] U.-S. Khoo, H. Ozcelik, A. N. Cheung, L. W. Chow, H. Y. Ngan, S. J. Done, A. Liang, V. W. Chan, G. K. Au, W.-F. Ng, *et al.*, "Somatic mutations in the brca1 gene in chinese sporadic breast and ovarian cancer", *Oncogene*, vol. 18, no. 32, p. 4643, Aug. 1999.

[157] R. Scully, J. Chen, A. Plug, Y. Xiao, D. Weaver, J. Feunteun, T. Ashley, and D. M. Livingston, "Association of brca1 with rad51 in mitotic and meiotic cells", *Cell*, vol. 88, no. 2, pp. 265–275, Jan. 1997.

[158] Q. Zhong, C.-F. Chen, S. Li, Y. Chen, C.-C. Wang, J. Xiao, P.-L. Chen, Z. D. Sharp, and W.-H. Lee, "Association of brca1 with the hrad50-hmre11-p95 complex and the dna damage response", *science*, vol. 285, no. 5428, pp. 747–750, Jul. 1999.

[159] A. Folias, M. Matkovic, D. Bruun, S. Reid, J. Hejna, M. Grompe, A. D'andrea, and R. Moses, "Brca1 interacts directly with the fanconi anemia protein fanca", *Human molecular genetics*, vol. 11, no. 21, pp. 2591–2597, Oct. 2002.

[160] P. Bork, K. Hofmann, P. Bucher, A. Neuwald, S. Altschul, and E. Koonin, "A superfamily of conserved domains in dna damage-responsive cell cycle checkpoint proteins.", *The FASEB Journal*, vol. 11, no. 1, pp. 68–76, Jan. 1997.

[161] I. A. Manke, D. M. Lowery, A. Nguyen, and M. B. Yaffe, "Brct repeats as phosphopeptide-binding modules involved in protein targeting", *Science*, vol. 302, no. 5645, pp. 636–639, Oct. 2003.

[162] X. Yu, C. C. S. Chini, M. He, G. Mer, and J. Chen, "The brct domain is a phospho-protein binding domain", *Science*, vol. 302, no. 5645, pp. 639–642, Oct. 2003.

[163] M. S. Chapman and I. M. Verma, "Transcriptional activation by brca1", *Nature*, vol. 382, no. 6593, p. 678, Aug. 1996.

[164] A. N. Monteiro, A. August, and H. Hanafusa, "Evidence for a transcriptional activation function of brca1 c-terminal region", *Proceedings of the National Academy of Sciences*, vol. 93, no. 24, pp. 13 595–13 599, Nov. 1996.

[165] L. Pellegrini, S. Y. David, T. Lo, S. Anand, M. Lee, T. L. Blundell, and A. R. Venkitaraman, "Insights into dna recombination from the structure of a rad51–brca2 complex", *Nature*, vol. 420, no. 6913, p. 287, Nov. 2002.

[166] M. Takata, S. Tachiiri, A. Fujimori, L. H. Thompson, Y. Miki, M. Hiraoka, S. Takeda, and M. Yamazoe, "Conserved domains in the chicken homologue of brca2", *Oncogene*, vol. 21, no. 7, p. 1130, Feb. 2002.

[167] H. Yang, P. D. Jeffrey, J. Miller, E. Kinnucan, Y. Sun, N. H. Thomä, N. Zheng, P.-L. Chen, W.-H. Lee, and N. P. Pavletich, "Brca2 function in dna binding and recombination from a brca2-dss1-ssdna structure", *Science*, vol. 297, no. 5588, pp. 1837–1848, Sep. 2002.

[168] P. Kerr and A. Ashworth, "New complexities for brca1 and brca2", *Current Biology*, vol. 11, no. 16, R668–R676, Aug. 2001.

[169] P.-L. Chen, C.-F. Chen, Y. Chen, J. Xiao, Z. D. Sharp, and W.-H. Lee, "The brc repeats in brca2 are critical for rad51 binding and resistance to methyl methanesulfonate treatment", *Proceedings of the National Academy of Sciences*, vol. 95, no. 9, pp. 5287–5292, Apr. 1998.

[170] S. Gretarsdottir, S. Thorlacius, R. Valgardsdottir, S. Gudlaugsdottir, S. Sigurdsson, M. Steinarsdottir, J. G. Jonasson, K. Anamthawat-Jonsson, and J. E. Eyfjörd, "Brca2 and p53 mutations in primary breast cancer in relation to genetic instability", *Cancer research*, vol. 58, no. 5, pp. 859–862, Mar. 1998.

[171] M. Tirkkonen, O. Johannsson, B. A. Agnarsson, H. Olsson, S. Ingvarsson, R. Karhu, M. Tanner, J. Isola, R. B. Barkardottir, Å. Borg, *et al.*, "Distinct somatic genetic changes associated with tumor progression in carriers of brca1 and brca2 germ-line mutations", *Cancer research*, vol. 57, no. 7, pp. 1222–1227, Apr. 1997.

[172] I. Callebaut and J.-P. Mornon, "From brca1 to rap1: A widespread brct module closely associated with dna repair", *FEBS letters*, vol. 400, no. 1, pp. 25–30, 1997.

[173] L.-C. Hsu and R. L. White, "Brca1 is associated with the centrosome during mitosis", *Proceedings of the National Academy of Sciences*, vol. 95, no. 22, pp. 12 983–12 988, Oct. 1998.

[174] X. Xu, Z. Weaver, S. P. Linke, C. Li, J. Gotay, X.-W. Wang, C. C. Harris, T. Ried, and C.-X. Deng, "Centrosome amplification and a defective g2–m cell cycle checkpoint induce genetic instability in brca1 exon 11 isoform–deficient cells", *Molecular cell*, vol. 3, no. 3, pp. 389–395, Mar. 1999.

[175] L. H. Castilla, F. J. Couch, M. R. Erdos, K. F. Hoskins, K. Calzone, J. E. Garber, J. Boyd, M. B. Lubin, M. L. Deshano, L. C. Brody, *et al.*, "Mutations in the brca1 gene in families with early-onset breast and ovarian cancer", *Nature genetics*, vol. 8, no. 4, p. 387, Dec. 1994.

[176] S. Gad, M. Klinger, V. Caux-Moncoutier, S. Pages-Berhouet, M. Gauthier-Villars, I. Coupier, A. Bensimon, A. Aurias, and D. Stoppa-Lyonnet, "Bar code screening on combed dna for large rearrangements of the brca1 and brca2 genes in french breast cancer families", *Journal of medical genetics*, vol. 39, no. 11, pp. 817–821, 2002.

[177] E. M. Rohlfs, N. Puget, M. L. Graham, B. L. Weber, J. E. Garber, C. Skrzynia, J. L. Halperin, G. M. Lenoir, L. M. Silverman, and S. Mazoyer, "An alu-mediated 7.1 kb deletion of brca1 exons 8 and 9 in breast and ovarian cancer families that results in alternative splicing of exon 10", *Genes, Chromosomes and Cancer*, vol. 28, no. 3, pp. 300–307, Jun. 2000.

[178] J. Simard, P. Tonin, F. Durocher, K. Morgan, J. Rommens, S. Gingras, C. Samson, J.-F. Leblanc, C. Belanger, F. Dion, *et al.*, "Common origins of brca1 mutations in canadian breast and ovarian cancer families", *Nature genetics*, vol. 8, no. 4, p. 392, 1994.

[179] C. Oddoux, J. P. Struewing, C. M. Clayton, S. Neuhausen, L. C. Brody, M. Kaback, B. Haas, L. Norton, P. Borgen, S. Jhanwar, *et al.*, "The carrier frequency of the brca2 6174delt mutation among ashkenazi jewish individuals is approximately 1%", *Nature genetics*, vol. 14, no. 2, p. 188, Oct. 1996.

[180] T. Wagner, R. Möslinger, C. Zielinski, O. Scheiner, and H. Breiteneder, "New austrian mutation in brca1 gene detected in three unrelated hboc families", *The Lancet*, vol. 347, no. 9010, p. 1263, 1996.

[181] S. L. Neuhausen, "Ethnic differences in cancer risk resulting from genetic variation", *Cancer: Interdisciplinary International Journal of the American Cancer Society*, vol. 86, no. S11, pp. 2575–2582, Nov. 1999.

[182] K. Claes, E. Machackova, M. De Vos, B. Poppe, A. De Paepe, and L. Messiaen, "Mutation analysis of the brca1 and brca2 genes in the belgian patient population and identification of a belgian founder mutation brca1 ivs5", *Disease markers*, vol. 15, no. 1-3, pp. 69–73, 1999.

[183] A. Liede, B. Cohen, D. Black, R. Davidson, A. Renwick, E. Hoodfar, O. Olopade, M. Micek, V. Anderson, R. De Mey, *et al.*, "Evidence of a founder brca1 mutation in scotland", *British journal of cancer*, vol. 82, no. 3, p. 705, Jan. 2000.

[184] L. Sarantaus, P. Huusko, H. Eerola, V. Launonen, P. Vehmanen, K. Rapakko, E. Gillanders, K. Syrjäkoski, T. Kainu, P. Vahteristo, *et al.*, "Multiple founder effects and geographical clustering of brca1 and brca2 families

in finland", *European Journal of Human Genetics*, vol. 8, no. 10, p. 757, Sep. 2000.

[185] A. Bergman, Z. Einbeigi, U. Olofsson, Z. Taib, A. Wallgren, P. Karlsson, J. WahlstroÈm, T. Martinsson, and M. Nordling, "The western swedish brca1 founder mutation 3171ins5; a 3.7 cm conserved haplotype of today is a reminiscence of a 1500-year-old mutation", *European Journal of Human Genetics*, vol. 9, no. 10, p. 787, Oct. 2001.

[186] M. Naeem, N. Khan, Z. Aman, A. Nasir, A. Samad, A. Khattak, *et al.*, "Pattern of breast cancer: Experience at lady reading hospital, peshawar", *J Ayub Med Coll Abbottabad*, vol. 20, no. 4, pp. 22–5, 2008.

[187] A. Amadou, P. Hainaut, and I. Romieu, "Role of obesity in the risk of breast cancer: Lessons from anthropometry", *Journal of oncology*, vol. 2013, 2013.

[188] S. Sohail and S. N. Alam, "Breast cancer in pakistan-awareness and early detection", *Journal of College of Physicians and Surgeons Pakistan*, vol. 7, pp. 711–712, 2007.

[189] F. Badar, S. Mahmood, R. Faraz, A. Yousaf, A. ul Quader, H. Asif, and A. Yousaf, "Epidemiology of breast cancer at the shaukat khanum memorial cancer hospital and research center, lahore, pakistan", *Journal of College of Physicians and Surgeons Pakistan*, vol. 25, no. 10, pp. 738–742, 2015.

[190] K. Masood, A. Masood, J. Zafar, A. Shahid, M. Kamran, S. Murad, M. Masood, Z. Alluddin, M. Riaz, N. Akhter, *et al.*, "Trends and analysis of cancer incidence for common male and female cancers in the population of punjab province of pakistan during 1984 to 2014", *Asian Pacific Journal of Cancer Prevention*, vol. 16, no. 13, pp. 5297–304, 2015.

[191] M. Ahmadian, M. Redzuan, Z. Emby, and A. A. Samah, "Women's community participation levels in community-based health programs regarding breast cancer prevention in metropolitan tehran, iran", *Asian Social Science*, vol. 6, no. 9, p. 12, 2010.

[192] Y. Bhurgri, N. Kayani, N. Faridi, S. Pervez, A. Usman, H. Bhurgri, J. Malik, I. Bashir, A. Bhurgri, S. H. Hasan, *et al.*, "Patho-epidemiology of breast cancer in karachi1995-1997'", *Asian Pacific Journal of Cancer Prevention*, vol. 8, no. 2, p. 215, 2007.

[193] M. A. Sharif, N. Mamoon, S. Mushtaq, and M. T. Khadim, "Morphological profile and association of her-2/neu with prognostic markers in breast carcinoma in northern pakistan", *J Coll Physicians Surg Pak*, vol. 19, no. 2, pp. 99–103, 2009.

[194] M. S. Siddiqui, N. Kayani, S. Sulaiman, A. Hussainy, S. Shah, and S. Muzaffar, "Breast carcinoma in pakistani females: A morphological study of 572 breast specimens.", *JPMA. The Journal of the Pakistan Medical Association*, vol. 50, no. 6, pp. 174–177, Jun. 2000.

[195] S. A. Bidgoli, R. Ahmadi, and M. D. Zavarhei, "Role of hormonal and environmental factors on early incidence of breast cancer in iran", *Science of the total environment*, vol. 408, no. 19, pp. 4056–4061, Sep. 2010.

[196] I. Jatoi and W. F. Anderson, "Qualitative age interactions in breast cancer studies: A mini-review", *Future oncology*, vol. 6, no. 11, pp. 1781–1788, Dec. 2010.

[197] A. Must, J. Spadano, E. H. Coakley, A. E. Field, G. Colditz, and W. H. Dietz, "The disease burden associated with overweight and obesity", *Jama*, vol. 282, no. 16, pp. 1523–1529, Oct. 1999.

[198] G. Berclaz, S. Li, K. Price, A. Coates, M. Castiglione-Gertsch, C.-M. Rudenstam, S. Holmberg, J. Lindtner, D. Erien, J. Collins, *et al.*, "Body mass index as a prognostic feature in operable breast cancer: The international breast cancer study group experience", *Annals of Oncology*, vol. 15, no. 6, pp. 875–884, Jun. 2004.

[199] V. Ozmen, B. Ozcinar, H. Karanlik, N. Cabioglu, M. Tukenmez, R. Disci, T. Ozmen, A. Igci, M. Muslumanoglu, M. Kecer, *et al.*, "Breast cancer risk factors in turkish women–a university hospital based nested case control study", *World Journal of Surgical Oncology*, vol. 7, no. 1, p. 37, Apr. 2009.

[200] G. M. Gilani, S. Kamal, and S. A. M. Gilani, "Risk factors for breast cancer for women in punjab, pakistan: Results from a case-control study", *Pakistan Journal of Statistics and Operation Research*, vol. 2, no. 1, pp. 17–26, Jan. 2006.

[201] X. Xia, W. Chen, J. Li, X. Chen, R. Rui, C. Liu, Y. Sun, L. Liu, J. Gong, and P. Yuan, "Body mass index and risk of breast cancer: A nonlinear dose-response meta-analysis of prospective studies", *Scientific reports*, vol. 4, p. 7480, Dec. 2014.

[202] W. H. Organization, *Issues of communication and risk: World health report 2002: From non-communicable diseases & mental health (nmh) communications*, 2002.

[203] H. Vainio, E. Weiderpass, and P. Kleihues, "Smoking cessation in cancer prevention", *Toxicology*, vol. 166, no. 1-2, pp. 47–52, Sep. 2001.

[204] G. A. Bray, "Medical consequences of obesity", *The Journal of Clinical Endocrinology & Metabolism*, vol. 89, no. 6, pp. 2583–2589, Jun. 2004.

[205] L. Bernstein, A. V. Patel, G. Ursin, J. Sullivan-Halley, M. F. Press, D. Deapen, J. A. Berlin, J. R. Daling, J. A. McDonald, S. A. Norman, *et al.*, "Lifetime recreational exercise activity and breast cancer risk among black women and white women", *Journal of the National Cancer Institute*, vol. 97, no. 22, pp. 1671–1679, Nov. 2005.

[206] V. B. Sheppard, K. Makambi, T. Taylor, S. F. Wallington, J. Sween, and L. Adams-Campbell, "Physical activity reduces breast cancer risk in african american women", *Ethnicity & disease*, vol. 21, no. 4, p. 406, Sep. 2011.

[207] C. Osborne, G. V. Ostir, X. Du, M. K. Peek, and J. S. Goodwin, "The influence of marital status on the stage at diagnosis, treatment, and survival of older women with breast cancer", *Breast cancer research and treatment*, vol. 93, no. 1, pp. 41–47, Sep. 2005.

[208] S. Floud, A. Balkwill, D. Canoy, F. L. Wright, G. K. Reeves, J. Green, V. Beral, and B. J. Cairns, "Marital status and ischemic heart disease incidence and mortality in women: A large prospective study", *BMC medicine*, vol. 12, no. 1, p. 42, Mar. 2014.

[209] A. A. Aizer, M.-H. Chen, E. P. McCarthy, M. L. Mendu, S. Koo, T. J. Wilhite, P. L. Graham, T. K. Choueiri, K. E. Hoffman, N. E. Martin, *et al.*, "Marital status and survival in patients with cancer", *Journal of clinical oncology*, vol. 31, no. 31, p. 3869, Nov. 2013.

[210] L. Bernstein, "Epidemiology of endocrine-related risk factors for breast cancer", *Journal of mammary gland biology and neoplasia*, vol. 7, no. 1, pp. 3–15, Nov. 2002.

[211] J. L. Kelsey, M. D. Gammon, and E. M. John, "Reproductive factors and breast cancer.", *Epidemiologic reviews*, vol. 15, no. 1, p. 36, 1993.

[212] M. Ewertz, S. W. Duffy, H.-O. Adami, G. Kvåle, E. Lund, O. Meirik, A. Mellemgaard, I. Soini, and H. Tulinius, "Age at first birth, parity and risk of breast cancer: A meta-analysis of 8 studies from the nordic countries", *International journal of cancer*, vol. 46, no. 4, pp. 597–603, Oct. 1990.

[213] D. Parkin, "15. cancers attributable to reproductive factors in the uk in 2010", *British journal of cancer*, vol. 105, no. S2, S73, Dec. 2011.

[214] J. Kotsopoulos, J. Lubinski, L. Salmena, H. T. Lynch, C. Kim-Sing, W. D. Foulkes, P. Ghadirian, S. L. Neuhausen, R. Demsky, N. Tung, *et al.*, "Breastfeeding and the risk of breast cancer in brca1 and brca2 mutation carriers", *Breast Cancer Research*, vol. 14, no. 2, R42, Mar. 2012.

[215] A. Hadjisavvas, M. A. Loizidou, N. Middleton, T. Michael, R. Papachristoforou, E. Kakouri, M. Daniel, P. Papadopoulos, S. Malas, Y. Marcou, *et al.*, "An investigation of breast cancer risk factors in cyprus: A case control study", *BMC cancer*, vol. 10, no. 1, p. 447, Aug. 2010.

[216] N. Elkum, T. Al-Tweigeri, D. Ajarim, A. Al-Zahrani, S. M. B. Amer, and A. Aboussekhra, "Obesity is a significant risk factor for breast cancer in arab women", *BMC cancer*, vol. 14, no. 1, p. 788, Feb. 2014.

[217] E. Barrett-Connor, "Hormones and the health of women: Past, present, and future keynote address", *Menopause*, vol. 9, no. 1, pp. 23–31, Jan. 2002.

[218] P. A. Marchbanks, K. M. Curtis, M. G. Mandel, H. G. Wilson, G. Jeng, S. G. Folger, J. A. McDonald, J. R. Daling, L. Bernstein, K. E. Malone, *et al.*, "Oral contraceptive formulation and risk of breast cancer", *Contraception*, vol. 85, no. 4, pp. 342–350, Apr. 2012.

[219] S. A. Silvera, A. B. Miller, and T. E. Rohan, "Oral contraceptive use and risk of breast cancer among women with a family history of breast cancer: A prospective cohort study", *Cancer causes & control*, vol. 16, no. 9, pp. 1059–1063, Nov. 2005.

[220] V. Dumeaux, E. Alsaker, and E. Lund, "Breast cancer and specific types of oral contraceptives: A large norwegian cohort study", *International journal of cancer*, vol. 105, no. 6, pp. 844–850, May 2003.

[221] M. Pike, M. Krailo, B. Henderson, A. Duke, and S. Roy, "Breast cancer in young women and use of oral contraceptives: Possible modifying effect of formulation and age at use", *The Lancet*, vol. 322, no. 8356, pp. 926–929, Oct. 1983.

[222] A. Lupulescu, "Estrogen use and cancer incidence: A review", *Cancer investigation*, vol. 13, no. 3, pp. 287–295, 1995.

[223] A. Surakasula, G. C. Nagarjunapu, and K. Raghavaiah, "A comparative study of pre-and post-menopausal breast cancer: Risk factors, presentation, characteristics and management", *Journal of research in pharmacy practice*, vol. 3, no. 1, p. 12, 2014.

[224] C. G. on Hormonal Factors in Breast Cancer *et al.*, "Menarche, menopause, and breast cancer risk: Individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies", *The lancet oncology*, vol. 13, no. 11, pp. 1141–1151, Nov. 2012.

[225] O. Kaidar-Person, G. Bar-Sela, and B. Person, "The two major epidemics of the twenty-first century: Obesity and cancer", *Obesity surgery*, vol. 21, no. 11, pp. 1792–1797, Nov. 2011.

[226] E. E. Calle, C. Rodriguez, K. Walker-Thurmond, and M. J. Thun, "Overweight, obesity, and mortality from cancer in a prospectively studied cohort of us adults", *New England Journal of Medicine*, vol. 348, no. 17, pp. 1625–1638, Apr. 2003.

[227] A. G. Renehan, M. Tyson, M. Egger, R. F. Heller, and M. Zwahlen, "Body-mass index and incidence of cancer: A systematic review and meta-analysis of prospective observational studies", *The Lancet*, vol. 371, no. 9612, pp. 569–578, Feb. 2008.

[228] K. Y. Wolin, K. Carson, and G. A. Colditz, "Obesity and cancer", *The oncologist*, vol. 15, no. 6, pp. 556–565, May 2010.

[229] G. K. Reeves, K. Pirie, V. Beral, J. Green, E. Spencer, and D. Bull, "Cancer incidence and mortality in relation to body mass index in the million women study: Cohort study", *Bmj*, vol. 335, no. 7630, p. 1134, Nov. 2007.

[230] L. H. Kushi, C. Doyle, M. McCullough, C. L. Rock, W. Demark-Wahnefried, E. V. Bandera, S. Gapstur, A. V. Patel, K. Andrews, T. Gansler, *et al.*, "American cancer society guidelines on nutrition and physical activity for cancer prevention: Reducing the risk of cancer with healthy food choices and physical activity", *CA: a cancer journal for clinicians*, vol. 62, no. 1, pp. 30–67, Jan. 2012.

[231] A. Trentham-Dietz, P. A. Newcomb, B. E. Storer, M. P. Longnecker, J. Baron, E. R. Greenberg, and W. C. Willett, "Body size and risk of breast cancer", *American journal of epidemiology*, vol. 145, no. 11, pp. 1011–1019, Jun. 1997.

[232] P. A. Van Den Brandt, D. Spiegelman, S.-S. Yaun, H.-O. Adami, L. Beeson, A. R. Folsom, G. Fraser, R. A. Goldbohm, S. Graham, L. Kushi, *et al.*, "Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk", *American journal of epidemiology*, vol. 152, no. 6, pp. 514–527, Sep. 2000.

[233] L. Vatten and S. Kvinnsland, "Body height and risk of breast cancer. a prospective study of 23,831 norwegian women", *British journal of cancer*, vol. 61, no. 6, p. 881, Jun. 1990.

[234] S. Cold, S. Hansen, K. Overvad, and C. Rose, "A woman's build and the risk of breast cancer", *European Journal of Cancer*, vol. 34, no. 8, pp. 1163–1174, Jul. 1998.

[235] M. Awatef, G. Olfa, M. Kacem, L. Sami, H. Makram, and B. A. Slim, "Association between body mass index and risk of breast cancer in tunisian women", *Annals of Saudi medicine*, vol. 31, no. 4, p. 393, 2011.

[236] X.-O. Shu, F. Jin, Q. Dai, J. R. Shi, J. D. Potter, L. A. Brinton, J. R. Hebert, Z. Ruan, Y.-T. Gao, and W. Zheng, "Association of body size and fat distribution with risk of breast cancer among chinese women", *International journal of cancer*, vol. 94, no. 3, pp. 449–455, Aug. 2001.

[237] L. M. Morimoto, E. White, Z. Chen, R. T. Chlebowski, J. Hays, L. Kuller, A. M. Lopez, J. Manson, K. L. Margolis, P. C. Muti, *et al.*, "Obesity, body size, and risk of postmenopausal breast cancer: The women's health initiative (united states)", *Cancer Causes & Control*, vol. 13, no. 8, pp. 741–751, Oct. 2002.

[238] A. Montazeri, J. Sadighi, F. Farzadi, F. Maftoon, M. Vahdaninia, M. Ansari, A. Sajadian, M. Ebrahimi, S. Haghighat, and I. Harirchi, "Weight, height, body mass index and risk of breast cancer in postmenopausal women: A case-control study", *BMC cancer*, vol. 8, no. 1, p. 278, Sep. 2008.

[239] D. Jung and S.-M. Lee, "Bmi and breast cancer in korean women: A meta-analysis", *Asian nursing research*, vol. 3, no. 1, pp. 31–40, Mar. 2009.

[240] B. Secretan, K. Straif, R. Baan, Y. Grosse, F. El Ghissassi, V. Bouvard, L. Benbrahim-Tallaa, N. Guha, C. Freeman, L. Galichet, *et al.*, "A review of human carcinogens—part e: Tobacco, areca nut, alcohol, coal smoke, and salted fish", *The lancet oncology*, vol. 10, no. 11, pp. 1033–1034, Nov. 2009.

[241] N. E. Collishaw, N. F. Boyd, K. P. Cantor, S. K. Hammond, K. C. Johnson, J. Millar, A. B. Miller, M. Miller, J. R. Palmer, A. G. Salmon, *et al.*, "Canadian expert panel on tobacco smoke and breast cancer risk", *Toronto, Canada: Ontario Tobacco Research Unit, OTRU Special Report Series*, 2009.

[242] T. W. Strine, J. M. Hootman, D. P. Chapman, C. A. Okoro, and L. Balluz, "Health-related quality of life, health risk behaviors, and disability among adults with pain-related activity difficulty", *American journal of public health*, vol. 95, no. 11, pp. 2042–2048, Nov. 2005.

[243] A. T. Kaczynski, S. R. Manske, R. C. Mannell, and K. Grewal, "Smoking and physical activity: A systematic review", *American journal of health behavior*, vol. 32, no. 1, pp. 93–110, 2008.

[244] R. R. Huxley, A. Ansary-Moghaddam, P. Clifton, S. Czernichow, C. L. Parr, and M. Woodward, "The impact of dietary and lifestyle risk factors on risk of colorectal cancer: A quantitative overview of the epidemiological evidence", *International journal of cancer*, vol. 125, no. 1, pp. 171–180, Jul. 2009.

[245] R. Baan, K. Straif, Y. Grosse, B. Secretan, F. El Ghissassi, V. Bouvard, A. Altieri, and V. Cogliano, "Carcinogenicity of alcoholic beverages", *The lancet oncology*, vol. 8, no. 4, pp. 292–293, Apr. 2007.

[246] A. Ikeda, I. Kawachi, H. Iso, M. Iwasaki, M. Inoue, and S. Tsugane, "Social support and cancer incidence and mortality: The jphc study cohort ii", *Cancer Causes & Control*, vol. 24, no. 5, pp. 847–860, May 2013.

[247] M. S. Rendall, M. M. Weden, M. M. Favreault, and H. Waldron, "The protective effect of marriage for survival: A review and update", *Demography*, vol. 48, no. 2, p. 481, Apr. 2011.

[248] L. Yang and K. H. Jacobsen, "A systematic review of the association between breastfeeding and breast cancer", *Journal of women's health*, vol. 17, no. 10, pp. 1635–1645, 2008.

[249] L. Lipworth, L. R. Bailey, and D. Trichopoulos, "History of breast-feeding in relation to breast cancer risk: A review of the epidemiologic literature", *Journal of the National Cancer Institute*, vol. 92, no. 4, pp. 302–312, Feb. 2000.

[250] S. Chang, J. R. Alderfer, L. Asmar, and A. U. Buzdar, "Inflammatory breast cancer survival: The role of obesity and menopausal status at diagnosis", *Breast cancer research and treatment*, vol. 64, no. 2, pp. 157–163, Nov. 2000.

[251] J. M. Petrelli, E. E. Calle, C. Rodriguez, and M. J. Thun, "Body mass index, height, and postmenopausal breast cancer mortality in a prospective cohort of us women", *Cancer Causes & Control*, vol. 13, no. 4, pp. 325–332, May 2002.

[252] J. J. Dignam, K. Wieand, K. A. Johnson, B. Fisher, L. Xu, and E. P. Mamounas, "Obesity, tamoxifen use, and outcomes in women with estrogen receptor–positive early-stage breast cancer", *Journal of the National Cancer Institute*, vol. 95, no. 19, pp. 1467–1476, Oct. 2003.

[253] S. Enger and L. Bernstein, "Exercise activity, body size and premenopausal breast cancer survival", *British journal of cancer*, vol. 90, no. 11, p. 2138, Jun. 2004.

[254] S. Kyogoku, T. Hirohata, S. Takeshita, Y. Nomura, T. Shigematsu, and A. Horie, "Survival of breast-cancer patients and body size indicators", *International journal of cancer*, vol. 46, no. 5, pp. 824–831, Nov. 1990.

[255] M.-H. Tao, X.-O. Shu, Z. X. Ruan, Y.-T. Gao, and W. Zheng, "Association of overweight with breast cancer survival", *American journal of epidemiology*, vol. 163, no. 2, pp. 101–107, Dec. 2005.

Appendix-I

DNA Extraction Protocol used in the Present Study

Day 1

1. All the samples were arranged in order according to their ID's and were transferred to labeled (50 ml) falcon tubes.
2. Quantity of each sample was carefully recorded. Cell lysis buffer (KHCO3, NH4Cl and 0.5M EDTA) was added to each sample for the lysis of cells and removal of hemoglobin in a quantity equal to three times of the blood samples i.e. in 5ml blood, 15 ml of cell lysis buffer was added.
3. All the tubes were kept on ice for 30 minutes after addition of cell lysis buffer to stop the cellular metabolic activities followed by centrifugation (Eppendorf Refrigerated Centrifuge 5130) at 1200 rpm for 10 minutes at 4°C.
4. The supernatant was discarded and pallets were re-suspended followed by washing of the pallet with 15ml of cell lysis buffer in the same manner as mentioned above to entirely remove the hemoglobin and complete lysis of the remaining cells.
5. The supernatant was discarded and the pellet was re-suspended.
6. 4.75ml of Sodium Tris EDTA (STE) was added to each sample.
7. Then 250µl of 10% Sodium dodecyl sulfate (SDS) was added while vertexing the samples.
8. To the above mixture 10 µL of proteinase K enzyme (20mg/ml) (Fermentas, Lithuania) was added and samples were incubated at 55℃ for overnight in shaking water bath (Orbit Shaker Bath, Lab-Line, USA).

Day 2

1. The samples were extracted with 5ml (equal quantity) of equilibrated phenol (pH=8).
2. Agitated for 10 minutes and kept on ice or freezer for 10 minutes.
3. Then the samples were centrifuged (Eppendorf Refrigerated Centrifuge 5130) at 3200 rpm for 30 minutes at 4°C.
4. The aqueous phase was transferred to labeled (15ml) falcon tubes with cut tips method.
5. 5ml of chilled chloroform:isoamyl alcohol (24:1) was added to each sample and agitated for 10 minutes and kept on ice for 10 minutes.
6. Samples were centrifuged at 3200rpm for 30 minutes at 4℃ in refrigerated centrifuge (Eppendorf Refrigerated Centrifuge 5130).
7. The supernatant was separated with cut tip into the separate labeled 15ml centrifuge tubes.
8. To the isolated supernatant, 500µL of 10M ammonium acetate and 5ml of chilled Isopropanol was added and agitated until DNA precipitates as visible white threads.
9. The samples were then placed overnight at -20℃ (or for 15 minutes at -70℃).

Day 3

1. The stored samples were centrifuged at 3200 rpm for 1 hour at 4°C.
2. The supernatant was discarded and the DNA pellet was re-suspended by tapping the 15ml centrifuge tube.
3. 5ml of chilled 70% ethanol was added to the washed samples and again was centrifuged at 3200rpm for 40 minutes at 4°C.
4. The supernatant was again discarded and DNA was dried until the last drop disappeared.
5. After drying the pallets, TE buffer (10mM Tris; 1mM EDTA) was added to each sample according to the size of the pellet.
6. The complete DNA dissolving in TE buffer was achieved by incubating the sample tube in shaking water bath at 55℃ (Orbit Shaker Bath, Lab-Line, USA)DNA to expand freely.
7. Then whole solution was transferred to (1.5ml) tubes and was labeled as stock solution.
8. 20µl of each DNA was added to 80µl of ddH_2O to make 20% DNA dilution as working solution in a labeled (1.5ml) tubes.

Appendix-II

BRCA1 Gene Primers

Primer ID	Sequence	Length	Product Size
BRCA1_Ex1_F	ATGAAGTTGTCATTTTATAAACCTTTT	27	236
BRCA1_Ex1_R	GGTCAATTCTGTTCATTTGCAT	22	
BRCA1_Ex2_F	TTGAGGCCTTATGTTGACTCAG	22	250
BRCA1_Ex2_R	AGGTGTTTCCTGGGTTATGAA	21	
BRCA1_Ex3_F	CATGGCTATTTGCCTTTTGA	20	250
BRCA1_Ex3_R	GAATGGTTTTATAGGAACGCTATG	24	
BRCA1_Ex4_F	CATGTCTTTTCTTATTTTAGTGTCCTT	27	250
BRCA1_Ex4_R	CGTCATAGAAAGTAATTGTGCAAAC	25	
BRCA1_Ex5_F	GGGTTTCTCTTGGTTTCTTTGA	22	389
BRCA1_Ex5_R	AAAATTAGCCTGGCATGGTG	20	
BRCA1_Ex6_F	TGGTGTCAAGTTTCTCTTCAGG	22	300
BRCA1_Ex6_R	CACTTCCCAAAGCTGCCTAC	20	
BRCA1_Ex7_F	CTGCCACAGTAGATGCTCAG	20	250
BRCA1_Ex7_R	AAAAGAGAGAAACATCAATCCTTAAT	26	
BRCA1_Ex8_F	TGATCTTGGTCATTTGACAGTTCT	24	246
BRCA1_Ex8_R	AAGGTCCCAAATGGTCTTCA	20	
BRCA1_Ex10_F	TGAGAAGAAAAAGACACAGCAAG	23	250
BRCA1_Ex10_R	GCAAACCTAAGAATGTGGGATA	22	
BRCA1_Ex11_F	GGTGATTTCAATTCCTGTGC	20	400
BRCA1_Ex11_R	TTACCCATGTGCTGAGCAAG	20	
BRCA1_Ex12_F	TGTCTGTTGCATTGCTTGTG	20	392
BRCA1_Ex12_R	AAAAACTGGAGAAAGTATGGTGAAA	25	

Primer ID	Sequence	Length	Product Size
BRCA1_Ex13_F	GCTGCCCAGCAAGTATGATT	20	340
BRCA1_Ex13_R	GGTTAACCAGAATATCTTTATGTAGGA	27	
BRCA1_Ex14_F	AATTCTTAACAGAGACCAGAACTTTG	26	498
BRCA1_Ex14_R	TGTTTTCTAGATTTCTTCCTCTAGGTT	27	
BRCA1_Ex15_F	TGTAGAACGTGCAGGATTGC	20	298
BRCA1_Ex15_R	CAAAGTGCTGCGATTACAGG	20	
BRCA1_Ex16_F	TCTTTAGCTTCTTAGGACAGCACTT	25	250
BRCA1_Ex16_R	CTCAGCATCAGCAAAAACCTT	21	
BRCA1_Ex17_F	AGGGAAGGACCTCTCCTCTG	20	250
BRCA1_Ex17_R	GGTGCATTGATGGAAGGAAG	20	
BRCA1_Ex18_F	TGTCTGCTCCACTTCCATTG	20	242
BRCA1_Ex18_R	AAATGAAGCGGCCCATCT	18	
BRCA1_Ex19_F	AAGCTCTTCCTTTTTGAAAGTCTG	24	249
BRCA1_Ex19_R	CCATCGTGGGATCTTGCTTA	20	
BRCA1_Ex20_F	TTCATCCGGAGAGTGTAGGG	20	231
BRCA1_Ex20_R	ACTGACAGGTGCCAGTCTTG	20	
BRCA1_Ex21_F	CAGAGCAAGACCCTGTCTCA	20	246
BRCA1_Ex21_R	CTCAAGCACCAGGTAATGAGTG	22	
BRCA1_Ex22_F	AGGACCCTGGAGTCGATTG	19	275
BRCA1_Ex22_R	GGCCTGGAAAGGCCACTT	18	

Total 24 exons among which 22 are coding and given here

BRCA2 Gene Primers

Primer ID	Sequence	Length	Product Size
BRCA2_Ex1_F	AATGCATCCCTGTGTAAGTGC	21	224
BRCA2_Ex1_R	TGGGTTTTTAGCAAGCATTTTT	22	
BRCA2_Ex2_F	TGATCTTTAACTGTTCTGGGTCAC	24	399
BRCA2_Ex2_R	GCTAAGATTTTAACACAGGTTTGC	24	
BRCA2_Ex3_F	AAACACTTCCAAAGAATGCAAA	22	295
BRCA2_Ex3_R	TCTACCAGGCTCTTAGCCAAA	21	
BRCA2_Ex4_F	CCAACAATTTATATGAATGAGAATCTT	27	250
BRCA2_Ex4_R	CATACCACTGGGGGTAAAAA	20	
BRCA2_Ex5_F	AACACCACAAAGAGATAAGTCAGG	24	232
BRCA2_Ex5_R	TCTCAGGGCAAAGGTATAACG	21	
BRCA2_Ex6_F	GCGTTATACCTTTGCCCTGA	20	483
BRCA2_Ex6_R	GCTTGACACCACTGGACTACC	21	
BRCA2_Ex7_F	TGTGCTTTTTGATGTCTGACAAA	23	250
BRCA2_Ex7_R	TCTCAAAGGCTTAGATAAATTACAGA	26	
BRCA2_Ex8_F	GAAATCACCAAAAGTGAAACCA	22	300
BRCA2_Ex8_R	ACGGGTGACAGAGCAAGACT	20	
BRCA2_Ex9.1_F	TGTTTCTATGAGAAAGGTTGTGAGA	25	477
BRCA2_Ex9.1_R	ACCATTCACAGGCCAAAGAC	20	
BRCA2_Ex9.2_F	GCCCTTTGAGAGTGGAAGTG	20	495
BRCA2_Ex9.2_R	TCAATTTCAGAGGCTTCAGTTTC	23	
BRCA2_Ex9.3_F	TCACCTAAAGAGACTTTCAATGC	23	474
BRCA2_Ex9.3_R	GATATGAAGATTTTAAAAAGCAGAAAA	27	
BRCA2_Ex10.1_F	CACTGTGCCCAAACACTACC	20	496
BRCA2_Ex10.1_R	AAGAGTGCTGGCATTTTCAT	20	
BRCA2_Ex10.2_F	GCAGCATGTCACCCAGTACA	20	487
BRCA2_Ex10.2_R	TTTCATTAGCTACTTGGAAGACAAAA	26	
BRCA2_Ex10.3_F	TTTCAAAAATAACTGTCAATCCAGA	25	457
BRCA2_Ex10.3_R	TGTTCAGAGAGCTTGATTTCCTT	23	
BRCA2_Ex10.4_F	GGCAGGACTCTTAGGTCCAA	20	491
BRCA2_Ex10.4_R	AATCGATGGGGCATTCATTA	20	
BRCA2_Ex11_F	TGGTCAAAACAGAACAAAAATG	22	392
BRCA2_Ex11_R	CAGCACTTTGAGAGGCAGGT	20	
BRCA2_Ex12_F	TTGAGCATCTGTTACATTCACTG	23	362

Primer ID	Sequence	Length	Product Size
BRCA2_Ex12_R	TGAACAGCACTATAAAATACTACCAAA	27	
BRCA2_Ex13.1_F	ATGAGGGTCTGCAACAAAGG	20	379
BRCA2_Ex13.1_R	GCTTTTGTCTGTTTTCCTCCA	21	
BRCA2_Ex13.2_F	ACAGGCAGACCAACCAAAGT	20	325
BRCA2_Ex13.2_R	GGGGAAAACCATCAGGACAT	20	
BRCA2_Ex14_F	ATTACAGGCGTGAGCCACTG	20	391
BRCA2_Ex14_R	TCATTCATCCATTCCTGCAC	20	
BRCA2_Ex15_F	TTTGGTAAATTCAGTTTTGGTTTG	24	379
BRCA2_Ex15_R	GCCAACTTTTTAGTTCGAGAGA	22	
BRCA2_Ex16_F	AATGATCTTGAACAATGTAGTTTTTG	26	384
BRCA2_Ex16_R	CACTGACAACTGGCTTGTGC	20	
BRCA2_Ex17.1_F	TTTTATTCTCAGTTATTCAGTGACTTG	27	400
BRCA2_Ex17.1_R	CAGGAGAGCCCACCAGTTCT	20	
BRCA2_Ex17.2_F	TCCTCCCCTCTTAGCTGTCTT	21	243
BRCA2_Ex17.2_R	ACATCTAAGAAATTGAGCATCCTT	24	
BRCA2_Ex18_F	GGCAGTTCTAGAAGAATGAAAACTC	25	382
BRCA2_Ex18_R	GCTGCAGTGAACCAAGATCA	20	
BRCA2_Ex19_F	TGCCTGGCCTGATACAATTA	20	396
BRCA2_Ex19_R	TGTCCCTTGTTGCTATTCTTTG	22	
BRCA2_Ex20_F	TCTCCCTTCTTTGGGTGTTTT	21	299
BRCA2_Ex20_R	TCCTGTGATGGCCAGAGAGT	20	
BRCA2_Ex21_F	AACCACACCCTTAAGATGAGC	21	455
BRCA2_Ex21_R	GGGCATTAGTAGTGGATTTTGC	22	
BRCA2_Ex22_F	TCCACTACTAATGCCCACAAA	21	371
BRCA2_Ex22_R	CAGAAAACAAAACAAAAATTCAACA	25	
BRCA2_Ex23_F	TGTAATTTTTCAGTTTTGATAAGTGC	26	300
BRCA2_Ex23_R	AGCTCCAACTAATCATAAGAGATTTT	26	
BRCA2_Ex24_F	GAGTTTCCTTTCTTGCATCTTAAA	24	396
BRCA2_Ex24_R	AAGCTATTTCCTTGATACTGGACTG	25	
BRCA2_Ex25_F	GGTCCAAACTTTTCATTTCTGC	22	381
BRCA2_Ex25_R	CAGGAGCCACATAACAACCA	20	
BRCA2_Ex26.1_F	TGTGTGTAATATTTGCGTGCTT	22	497
BRCA2_Ex26.1_R	AATGCAAGTTCTTCGTCAGC	20	
BRCA2_Ex26.2_F	CATTTCAGCCACCAAGGAGT	20	480
BRCA2_Ex26.2_R	TTTCTTTTCTCATTGTGCAACATA	24	

Appendix-III

Proforma for Breast Cancer Sample Collection

BIOINFORMATICS AND EXPERIMENTAL ANALYSIS OF THE GENETIC AND NON-GENETIC BASIS OF BREAST CANCER IN PAKISTANI POPULATION

Department of Biosciences, Capital University of Science & Technology, Islamabad, Pakistan

You are invited to take part in this research study. This form tells you why this research study is being done, what will happen in the research study, and possible risks and benefits to you. If there is anything you do not understand, please ask questions. Then you can decide if you want to join this study or not. If you have read this form and have decided to participate in this project, please understand your participation is voluntary.

Your identity and personal data will NOT be known to any personnel other than the investigators and will not be disclosed in any published and written material resulting from the study. It is also to be noted that, you will not be paid to participate in this study.

Hospital ID: _____	Deptt. ID: _____	Sample No. _____
Status: _____	Sample Type: _____	Date: _____

1. PATIENTS DETAILS:

Sex: _____ DOB: _____ Religion: _____

Ethnicity/Caste: _____ Weight: _____ Height: _____

Marital status: _____ No. of Children: _____ Breast feeding: _____

BP: _____ Smoking: _____ Contact No: _____

Any other: _____

Address: _____

2. DISEASE HISTORY:

Age of disease onset: _____ Symptoms: _____

Treatment details: _____

Side effects: _____

Previous diagnosis: _____

Have you ever had any type of cancer? _____ Surgery status: _____

Do you walk (or do other moderate activity) for at least 30 min on most days:_____

3. FAMILY HISTORY:

Do you have any family member who has had any type of cancer? _____

If yes, please Specify: _____

Type of Cancer diagnosed in the family member: _____

4. DISEASE COMPLICATIONS:

- Breast Affected: _____
- Swelling of all or part of the breast: _____
- Skin irritation or dimpling: _____
- Breast/armpit pain: _____
- Nipples discharge other than milk: _____
- A lump in the under arm area: _____
- Disease Stage: _____

- Menopause status: _____
- Weight loss: _____
- Any other complication: _____

INFORMED CONSENT:

Since information about you and your health is personal and private, it generally cannot be used in this research study without your written authorization. If you sign this form, it will provide that authorization. Please read it carefully before signing it. I am donating my blood/tissue sample for research purposes only and not for commercial purposes.

Signature: _____