MUTATIONAL ANALYSIS OF MITOCHONDRIAL AND GLUTATHIONE S-TRANSFERASE GENES ASSOCIATED WITH BREAST CANCER RISK AMONG MIZO POPULATION

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

LALHMACHHUANI
MZU REGISTRATION: 235 of 2014
Ph.D. REGISTRATION: MZU/Ph.D./635 of 23.05.2014

DEPARTMENT OF ZOOLOGY
SCHOOL OF LIFE SCIENCES
JUNE, 2023

MUTATIONAL ANALYSIS OF MITOCHONDRIAL AND GLUTATHIONE S-TRANSFERASE GENES ASSOCIATED WITH BREAST CANCER RISK AMONG MIZO POPULATION

BY

Lalhmachhuani

Department of Zoology

Prof. G. Gurusubramanian

Submitted

In partial fulfillment of the requirement of the Degree of

Doctor of Philosophy in Zoology of Mizoram University, Aizawl

To the Ones who inspired me to walk the extra mile

Nothing in life is to be feared,

It is only to be understood

Marie Curie

Department of Zoology
School of Life Sciences
MIZORAM UNIVERSITY
(*A Central University*)

Dr G Gurusubramanian Ph.D
Professor

Phone: 9862399411 (M)

CERTIFICATE

I certify that the thesis entitled "**Mutational analysis of mitochondrial and glutathione s-transferase genes associated with breast cancer risk among Mizo population**" submitted to Mizoram University for the award of the degree of Doctor of Philosophy in Zoology by **Lalhmachhuani**, a record of original research work carried out during the period of 2014-2023 under my guidance and supervision, and that this work has not formed the basis for the award of any degree, diploma, associateship, fellowship or other titles in this university or any other university or institution of higher learning.

(Prof. G. GURUSUBRAMANIAN)

Department of Zoology

Mizoram University

MIZORAM UNIVERSITY

AIZAWL-796004

June, 2023

DECLARATION

I, Lalhmachhuani, hereby declare that the subject matter of this thesis is the record of work done by me, that the contents of this thesis did not form basis of the award of any previous degree to me or to the best of my knowledge to anybody else, and that the thesis has not been submitted by me for any research degree in any other University/Institute.

This is being submitted to the Mizoram University for the degree of Doctor of Philosophy in Zoology.

(LALHMACHHUANI)

(HEAD) **(SUPERVISOR)**

ACKNOWLEDGEMENT

This thesis even though I bear the name would not have been possible without the contributions of many which I may even miss a few. I would like to offer my sincere gratitude to:

Prof G Gurusubramanian, my invaluable guide for not giving up on me

Dr Jeremy L Pautu for all the challenges, who with a critical eye of a physician read the manuscript and helped me sort out some of my own misconceptions

Mizoram State Cancer Institute and Department of Pathology, Aizawl Civil Hospital for their valuable data; the Institutional Ethics Committee for the permission

Prof N Senthil Kumar for selecting the genes and guidance

Dr James LT Thanga for clearing all my doubts regarding SPSS and statistics, for the guidance and encouragement. To his lovely family for their hospitality and friendship

Dr Esther Lalhmingliani for all the support and encouragements

Department of Zoology for allowing me to do this work; and Prof Ram Pal Vadhera and the University

Vala, Sunaina, Judith, Puipuii, Lalmami, Maii and Lawmi for the sampling; Opa for the layout and Tetei for the tree

The Volunteers who participated in this study

Lastly, I offer my regards and blessings to all, especially friends and family who supported me in any respect until the completion of this work.

Dated :

Place : Aizawl **(LALHMACHHUANI)**

CONTENTS

Certificate from the Supervisor

Declaration

Acknowledgement

Table of Contents

List of Tables

List of Figures

	Page No.
I. INTRODUCTION	1 - 12
1.1. Biology of the Breast	4 - 5
1.2. Types of Breast Cancer	5 - 6
1.3. Epidemiology	7
1.4. Mitochondrial DNA	8 - 10
1.4.1. Displacement Loop	9
1.4.2. Cytochrome C Oxidase	9 - 10
1.5. Glutathione S-Transferase	10 - 12
II. REVIEW OF LITERATURE	13 - 46
2.1. Brief History of Breast Cancer	14 - 16
2.2. Epidemiology	16 - 40

2.2.1. Reproductive Factors	**19 - 24**
2.2.1.1. Age at Menarche and Menopause	**19 - 20**
2.2.1.2. Childbearing	**20 - 21**
2.2.1.3. Parity	**21 - 23**
2.2.1.4. Breastfeeding	**23 - 24**
2.2.2. Family history and Medical history	**24 - 26**
2.2.3. Lifestyle Factors	**26 - 33**
2.2.3.1. Anthropometric Factors	**26 - 27**
2.2.3.2. Physical Activity	**28 - 29**
2.2.3.3. Sleep Cycle and Circadian Disruption	**29**
2.2.3.4. Alcohol	**29 - 31**
2.2.3.5. Tobacco	**31 - 33**
2.2.3.6. Betel Quid	**33**
2.2.4. Diet and Nutrition	**34 -35**
2.2.5. Exogenous Hormones - Oral Contraceptives and Postmenopausal Hormone	**35 - 36**
2.2.6. Environmental and Occupational Exposures	**36 - 40**
2.2.6.1. Environmental Pollutants	**37 - 38**
2.2.6.2. Radiation – Ionizing and Non-Ionizing	**38 - 39**
2.2.6.3. Trace Elements and Heavy Metals	**39 - 40**
2.3. Mitochondrial DNA	**40 - 43**

2.3.1. Displacement Loop 41 - 42

2.3.2. Cytochrome C Oxidase 42 - 43

2.4. Glutathione S-transferases 43 - 46

2.4.1. Glutathione S-transferase Mu 44 - 45

2.4.2. Glutathione S-transferase Theta 45 - 46

III. OBJECTIVES 47 - 48

IV. MATERIALS AND METHODS 49 - 61

4.1. Data Source and Sample Collection 50

4.2. Questionnaire 50 - 52

4.3. Anthropometric and Physical Activity Measurement 52 - 53

4.4. Molecular Study 54 - 59

4.4.1. DNA Isolation 54 - 55

4.4.2. Amplification of Gene 55 - 58

4.5. DNA Analysis 58 - 59

4.6. Statistical Methods 59 - 61

V. RESULTS 62 - 111

5.1. Demographic Characteristics 63 - 66

5.2. Comparison of Variables 66 - 72

5.2.1. Demography 66

5.2.2. Diet and Nutrition 66

5.2.3. Behavioral Habits 69 - 70

5.2.4. Environmental Exposure and Medical History 70

5.2.5. Menstrual History 72

5.2.6. Reproductive History 72

5.3. Factors Related to Breast Cancer on Univariate Analysis 72

5.4. Factors Related to Breast Cancer on Multivariate Analysis 74 - 75

5.5. Comparison of Breast Cancer Cases Based on Survival Status 75 - 76

5.6. Factors Related to Breast Cancer on Cox-regression Model 77

5.7. Comparison of Groups Based on Menopausal Status 78 - 83

5.8. History of Breast and Other Cancers 83

5.9. Molecular Study 84 - 111

5.9.1. Comparison of Cases and Controls Groups 84

5.9.2. Association of Selected Variables 86

5.9.3. Glutathione S-transferase 90 - 97

5.9.4. Mitochondrial DNA 97 - 111

VI. DISCUSSION 112 - 127

6.1. Demographic Characteristics 114

6.2. Dietary Habits 114 - 116

6.3. Behavioral Habits **116 - 118**

6.4. Environmental Exposure and Medical History **118 - 119**

6.5. Menstrual History **120**

6.6. Reproductive History **120 - 121**

6.7. Association of Risk Factors with Glutathione S-transferase **122**

6.8. Association of Risk Factors with Mitochondrial DNA **123 - 125**

VII. SUMMARY **128 - 129**

VIII. APPENDIX **130 - 135**

1. **List of Acronyms** **130 - 131**
2. **Annexure 1** **132**
3. **Annexure 2** **133**
4. **Annexure 3** **134 - 135**

IX. REFERENCES **136 - 196**

BRIEF BIO-DATA OF THE CANDIDATE **197**

PARTICULARS OF THE CANDIDATE **198**

LIST OF TABLES

Table No.	Description	Page No.
1	Interview questionnaires	52
2	Physical activity of 24-hour duration	53
3	List of primers and sequence	55
4	Thermal profile for D-Loop	56
5	Thermal profile for CO1	56
6	Thermal profile for multiplex PCR	58
7	Age distribution of cases (1998-2014)	63
8	Breast cancer incidence (1998-2014) districtwise	64
9	Number of cases registered and interviewed	64
10	Comparison of mean and standard deviation between groups	65
11	Tumor histology and molecular sub-types of breast cancer cases	66
12	Comparison of demographic characteristics between groups	67
13	Comparison of dietary habits between groups	68
14	Comparison of behavioral habits between groups	69
15	Comparison of environmental exposure and chronic diseases between groups	70
16	Comparison of menstrual characteristics between groups	71
17	Comparison of reproductive characteristics between groups	73
18	Factors related to breast cancer between groups on univariate analysis	74
19	Factors related to breast cancer between groups on multivariate analysis	75
20	Comparison of breast cancer cases on survival status	76 - 77
21	Factors related to breast cancer on Cox-regression model	78
22	Comparison of selected characteristics of premenopausal groups	78 - 79
23	Univariate analysis of selected factors on premenopausal groups	80
24	Comparison of selected characteristics of postmenopausal groups	81 - 82
25	Univariate analysis of selected factors on postmenopausal groups	83
26	Cases with number of I, II and III degree relatives having cancer	83
27	Comparison of selected variables for molecular analysis	84
28	Comparison of dietary and behavioral characteristics for molecular analysis	85 - 86
29	Comparison of reproductive characteristics for molecular analysis	87
30	Selected variables of the subjects with adjusted OR for age and residence	88 - 89

31	Comparison of GSTT1 and GSTM1 between groups	91
32	Association of GSTT1 and GSTM1 genotype on menopausal status	93
33	Association of GSTT1/GSTM1 genotype on menopausal status	93
34	Association of GSTT1 with selected variables	94 - 95
35	Association of GSTM1 with selected variables	96 - 97
36	Comparison of variations found in D-loop region of groups	98 - 99
37	Comparison of variations found in CO1 region of groups	100
38	Association of haplogroups with study groups	101
39	Comparison of mitochondrial variations with significance	107
40	Association of mitochondrial variation at locus 146, 152 and 248 with breast cancer on menopausal status	107
41	Association of mitochondrial variation at locus 5945, 6392 and 6617 with breast cancer on menopausal status	108
42	Association of mitochondrial variation at locus 16304, 16311 and 16519 with breast cancer on menopausal status	108
43	Association of mitochondrial variations at locus 146, 152 and 248 with breast cancer risk	109
44	Association of mitochondrial variations at locus 5945, 6392 and 6617 with breast cancer risk	110
45	Association of mitochondrial variations at locus 16304, 16311 and 16519 with breast cancer risk	111

LIST OF FIGURES

Figure No.	Description	Page No.
1	Overview of Carcinogenesis	3
2	Anatomy of female breast	4
3	Types of breast cancer	6
4	Human mitochondrial DNA	8
5	Structure of Human GST	10
6	Percentage of new cases in females, Globocan 2020	17
7	Percentage of new cases in females (Mizoram), NCRPR 2020	17
8	Schema showing how ROS may affect mitochondrial and nuclear DNA leading to breast carcinogenesis	41
9	Results of multiplex PCR for GSTM1 and GSTT1	91
10	Comparison of haplogroups found in case and control groups	100
11	Branching of haplogroups from ancestral root (L3)	101
12	Comparison of major tribes in the study	102
13	Phylogenetic tree based on CO1 sequences	103
14	Phylogenetic tree based on d-loop sequences	104

CHAPTER 1

INTRODUCTION

1. Introduction

The Mizos are a group of tribes inhabiting a landlocked area located on the north eastern part of India called Mizoram. It has an area of 21,081 square kilometers located at 21°56'N to 24°31'N latitude and 92°16'E to 93°26'E longitude. Mizoram shares its boundary on the northern side with Tripura, Assam and Manipur; and an international boundary with Myanmar in the east and south, and Bangladesh in the west. The Mizos are different from mainland India in their race and ethnicity and have a unique lifestyle, customs and dietary habits.

There is a saying that many women who develop breast cancer have no known risk factors other than simply being women. But this alone does not explain the 'why', the randomness and the occasional incidence in men. We have come a long way from the time when breast cancer was first described in the Edwin Smith Papyrus, an ancient Egyptian medical text as an 'ulcer' and untreatable (Diamandopoulos, 1996). We know now that there is a hereditary component to breast cancer and that it tends to cluster in certain families. Studies of migrants from low-risk Asian population migrating to high-risk USA provided the first solid evidence that environmental determinants were responsible for most of the observed international and inter-ethnic differences in breast cancer incidence (Ziegler *et al*., 1993). The relative contributions of pure genetic effects and of lifestyle remain unclear, but it is becoming more and more evident that genotypic inheritance and lifestyle are probably inseparably intertwined. The reason could be because the combination of genetic factors and lifestyle makes us who we are, and determining our individual susceptibility to that disease (Key *et al.*, 2001).

Breast cancer incidence rates are high in more developed countries, whereas rates in less developed countries are low but increasing. The volume on breast cancer research is huge, but there is no proper consensus on their findings, which could be due to differences in sample sizes, geography, race and ethnicity of study populations, local customs, lifestyle, culture, and health care conditions (Parkin *et al.*, 2005). Due to

variation in the geographical, racial and ethnic distribution of the disease, it is of utmost importance to identify risk factors that is specific for that population and whether the established factors in other parts of the world is applicable or relevant for that particular region. The purpose of the investigations in this thesis is to cast some light on which recognized environmental and lifestyle factors could be responsible for breast cancer incidence among the Mizo population. The other objective is to determine whether mutations in the mitochondrial and glutathione s-transferase genes might interact with these environmental and lifestyle factors in increasing breast cancer risk.

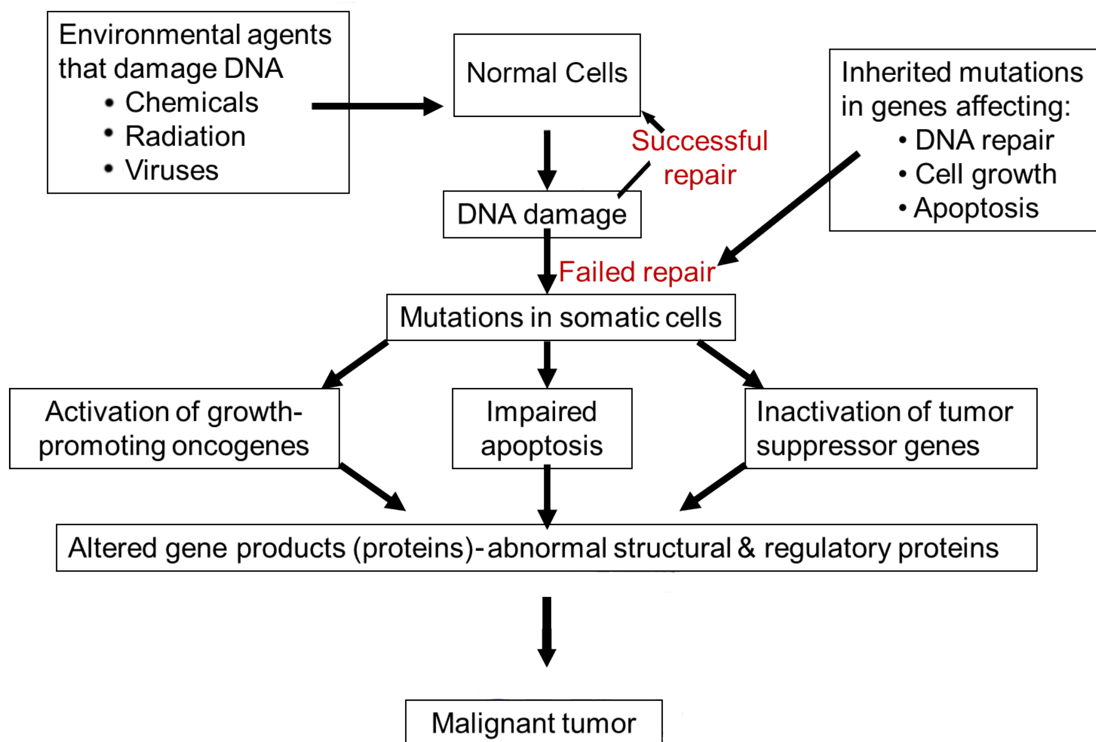


Figure 1. Overview of Carcinogenesis (https://sphweb.bumc.bu.edu).

1.1. Biology of the Breast

The basic component of a breast is alveoli made up of epithelial cuboidal cells embedded in myoepithelial cells. Each alveolus is lined by milk-secreting cells, the acini, which extract from the mammary blood supply the factors essential for milk formation. Each breast has 15 to 20 sections called lobes and within each lobe are many smaller lobules. Lobules are arranged in clusters like a bunch of grapes. Lobules unite together through a mesh work of ducts called lactiferous ducts. Lobules end in dozens of tiny bulbs that can produce milk. Thin tubes, called ducts, link all the lobes, lobules and bulbs. These ducts lead to the nipple in the center of a dark area of skin called the areola. Fat fills the spaces between lobules and ducts. There are no muscles in the breast, but muscles lie under each breast and cover the ribs. Branching of the ducts and alveolar growth occurs before puberty due to ovarian estrogen stimulation. Lobule formation in the female breast occurs within 1-2 years after the onset of the first menstrual period. Full differentiation of the mammary gland is a gradual process, which takes years to complete. It has also been observed that full maturation of secretory alveoli occurs at pregnancy, in which increased estrogen and progesterone levels cause further differentiation of ductal cells (Imagawa *et al*., 1990).

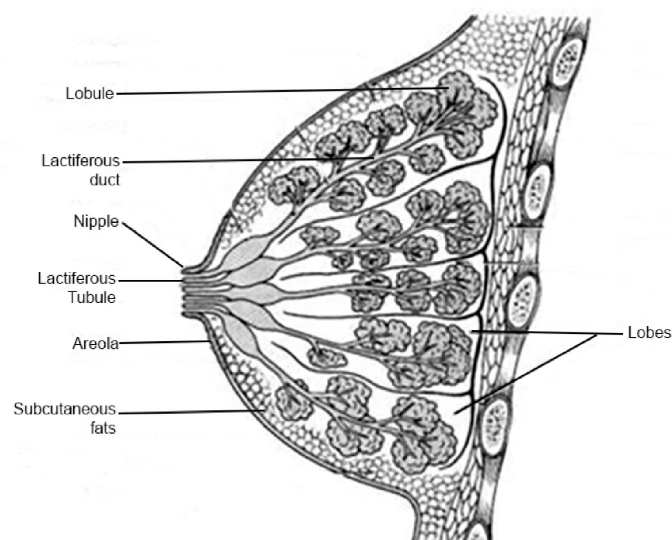


Figure 2. Anatomy of female breast (Verralls, 1993).

The breast of nulliparous women contains more undifferentiated structures while in premenopausal parous women the predominant structure is the most differentiated form. Full lobular differentiation only occurs in parous women, especially in those experiencing full term pregnancy at a young age. After menopause, the breast undergoes regression in both nulliparous and parous women manifested as an increase in the undifferentiated structures. Estrogens are a group of hormones that play a major role in promoting the proliferation of both normal and neoplastic breast epithelium. The amounts and types of estrogen present vary throughout life. The two types of estrogen that have been most closely studied in relation to breast cancer risk are estradiol and estrone sulfate. Estrone is the primary form after menopause which is produced mainly by fat cells, estradiol is the primary form during reproductive years made in ovaries, and estriol is the primary form during pregnancy. Estradiol acts locally on the mammary gland, stimulating DNA synthesis and promoting bud formation (Russo et al., 2000).

1.2. Types of Breast Cancer

All tumors arise from normal tissue and breast cancer is cancer that originates from breast tissue but the progression from normal breast tissue to invasive cancer is poorly understood. The most common type of breast cancer begins in the lining of the ducts called ductal carcinoma. The second most common is lobular carcinoma, which occurs in the lobes. Non-invasive breast cancer is called carcinoma in situ (CIS) and can arise from either ductal or lobular hyperplasia of epithelial cells (Wellings, 1980). Cancer that has progressed into surrounding tissue is called invasive breast cancer and usually has the ability to metastasize. Tumors are categorized according to type and size, histopathology, invasiveness, tumor stage and receptor expression. With improvement in molecular techniques, we now have a deeper understanding of diverse breast cancer types and how they differ (Simpson *et al*., 2005).

According to WHO, tumors are classified into six main types - ductal, lobular, mucinous, medullary, papillary and tubular carcinoma. Ductal and lobular tumors represent around 90-95% of all cases (Tavassoli and Devilee, 2003). Histological grade is often classified according to the Nottingham Grade classification which was introduced in the 1990s and includes three different parameters - tubule formation, nuclear pleomorphism and mitotic counts (Elston and Ellis, 1991). Tumor stage classification incorporates Tumor size (T), lymph Node status (N) and Metastasis (M) (usually shortened to TNM). The TNM system has been somewhat controversial but remains well used by clinicians (Benson *et al*., 2003; Cserni *et al*., 2018). Expression of different receptors, known to affect the prognostic and predictive values of therapy, is also used to characterize the tumors. They are classified according to expression of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2). The lack of any ER or PR on the tumor cells makes the prognosis worse whereas the lack of HER2 expression does not. If the tumor lacks all three receptors it is called triple negative, this feature often indicates a poor prognosis (Bauer *et al*., 2007; Parise *et al*., 2009).

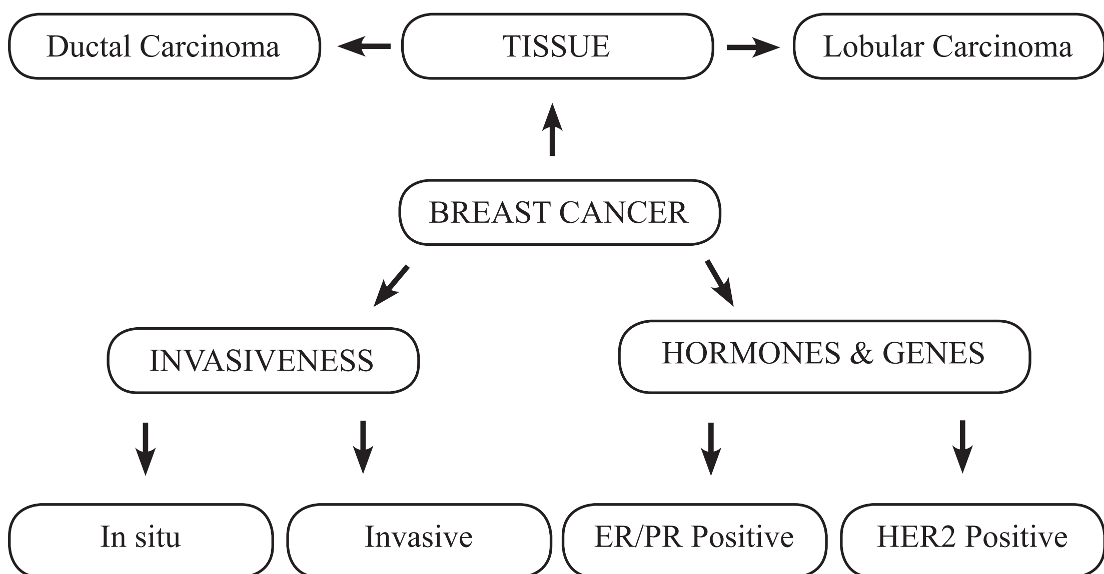


Figure 3. Types of breast cancer.

1.3. Epidemiology

Risk factor is anything that affects the chance of getting breast cancer. A statement by Jose Russo sums up the most appropriate definition of one's risk saying "Breast cancer is a complex disease caused by multiple environmental and lifestyle factors interacting with genetic susceptibility across the life span." There are important findings that establish certain agents as risk factors. These studies have shown that there are modifiable and non-modifiable factors that influence breast cancer risk. There are several controversial risk factors which are uncertain but suspected to cause breast cancer as well. Multiple occurrences of these risk factors in a person increase the likelihood of developing breast cancer.

According to WHO, breast cancer is the most frequent cancer among women, impacting 2.1 million women each year, and among women the cause for the greatest number of cancer-related deaths. According to GLOBOCAN 2012, among women, 25.2 percent of incident sites of cancer was the breast and had a substantially higher incidence (43.3 per 100 000) than any other cancer; the next highest incidence was of colorectal cancer (14.3 per 100 000). In 2018, it was estimated that 15 percent of all cancer deaths among women was due to breast cancer. Breast cancer rates used to be higher among women in more developed regions, but according to recent statistics, it is increasing globally in nearly every region. In India, roughly, one in four newly detected cancers in women is breast cancer. In 2018, 1,62,468 new cases and 87,090 deaths due to breast cancer were reported. The global distribution of cancer indicates marked, and sometimes extreme differences with respect to particular tumor type, which could be the key to understanding causation, and hence the development of preventive measures. In Mizoram, according to PBCR 2012-2014 report, breast cancer is the third most common cancer among females (13.5%), but we see rise in incidence and could lead to the most common cancer among females.

1.4. Mitochondrial DNA

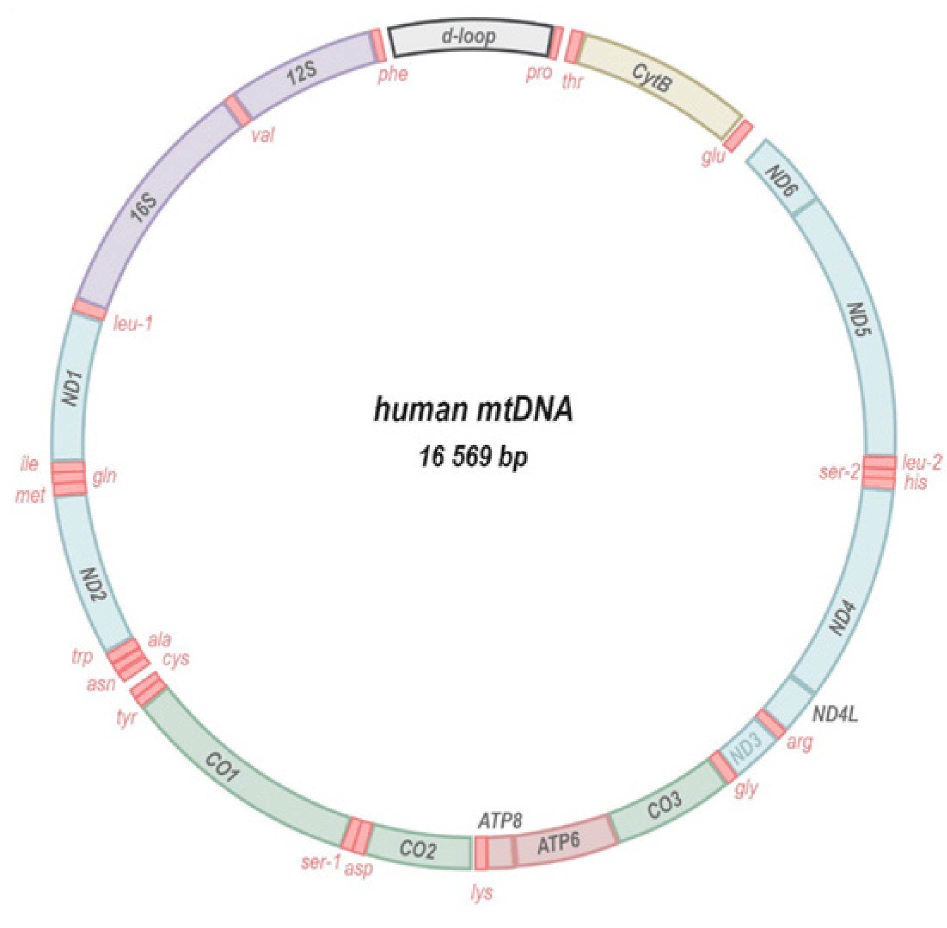


Figure 4. Human mitochondrial DNA (https://www.mdpi.com/2076-3921/8/9/392).

Mitochondria play a central role in oxidative metabolism and each cell contains thousands of copies of mitochondrial DNA (mtDNA). The human mitochondrial genome is a double stranded 16.6 kb circular DNA and is formed by a light strand and a heavy strand. The bright pink color bands denote tRNA and coding sequences are labeled by the respective amino acid code. Protein coding segments on mtDNA do not have introns and are transcribed by a single polycistronic mRNA from each strand. All protein coding sequences are marked with respective gene name abbreviations (Cyt b for Cytochrome b; ND for NADH dehydrogenase; CO for cytochrome c oxidase; and ATPase for ATP synthase). The two ribosomal (rRNAs) 12S and 16S RNAs locations are marked in lilac color. D-loop (grey color) region does not have coding sequences

(Singh, 1998; Taanman, 1999; Bianchi *et al*., 2001). Mutations in the mtDNA have been found in connection to various types of human cancer. Since the mtDNA encode several polypeptides of the respiratory-chain enzymes, mtDNA mutations often affect the function of oxidative phosphorylation. Some of the identified mutations cause amino-acid substitutions in the enzyme cytochrome c oxidase. Mitochondrial DNA is particularly susceptible to damage by environmental carcinogens because it contains no introns and lack histones, and is in close proximity to reactive oxygen species (ROS) produced through oxidative phosphorylation. Due to this, the mutation frequency in mtDNA is approximately tenfold greater than that in nuclear DNA (Johns, 1995; Grossman and Shoubridge, 1996).

1.4.1. Displacement Loop

The d-loop is a triple-stranded non-coding region of mtDNA, 1124 bp in size (np 16024-516) that contains cis-regulatory elements required for the replication and transcription of the mtDNA. All other mitochondrial proteins, including those involved in the replication, transcription and translation of mtDNA are nuclear encoded (Schatz, 1996). The d-loop region contains the leading strand for the origin of replication and a number of major promoters for transcription of the mitochondrial genome. Hence, it is possible that genetic variability in the d-loop region may affect the function of the respiration chain that is responsible for high reactive oxygen species levels and could contribute to nuclear genome damage and cancer initiation and progression. Moreover, respiratory chain alteration may cause a dysfunction in mitochondrion induced apoptosis (Ye *et al*., 2010).

1.4.2. Cytochrome C Oxidase

Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. It is the rate-limiting step of the mitochondrial electron transport chain (Villani *et al*., 1998) and represents a molecular switch that

induces apoptosis under energy stress conditions (Kadenbach *et al*., 2004). Defects in the mitochondrial genome and function are suspected to contribute to the development and progression of cancer (Ye *et al*., 2008). Several of the cancer associated mutations found in mtDNA result in structural modifications of cytochrome c oxidase (Namslauer and Brzezinksi, 2009). Of the 13 subunits that comprise cytochrome c oxidase, subunit I (np 5904-7445) is responsible for the control of apoptosis through phosphorylation or dephosphorylation events (Lee *et al*., 2001).

1.5. Glutathione S-Transferase

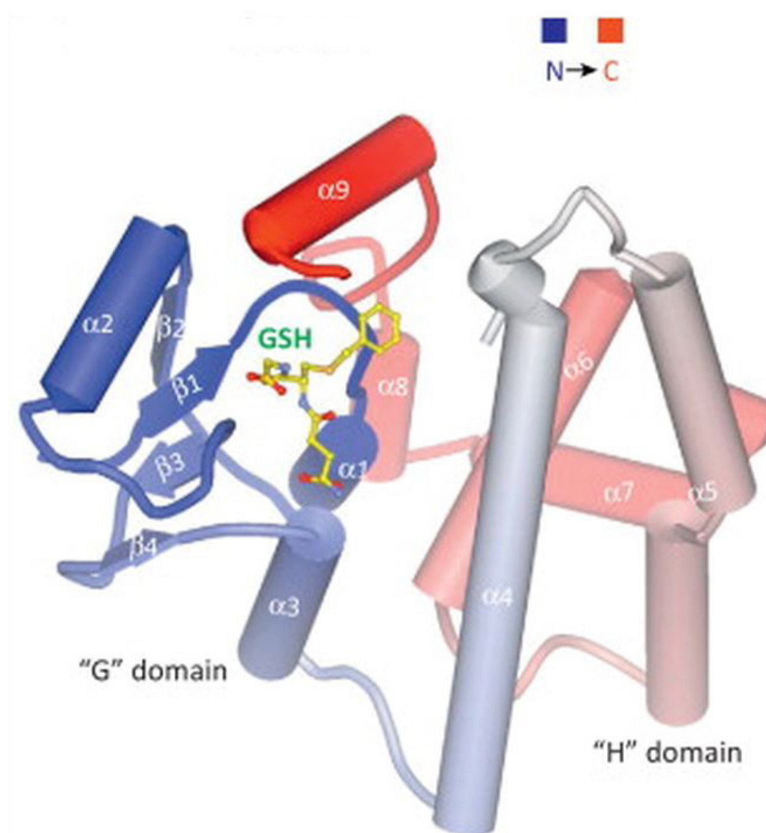


Figure 5. Structure of Human GST (Wu and Dong, 2012).

The glutathione s-transferases (GSTs) are complex superfamily of dimeric phase II metabolic enzymes involved in detoxification of a wide range of harmful chemicals, including environmental pollutants, carcinogens, mutagens and toxic products such as lipid hydrogen peroxides generated during oxidative stress. A single

GST unit consists of two subunits an N-terminal α/β-domain or G domain for binding glutathione (GSH) and an all-α-helical domain or H domain for binding hydrophobic substrates. The residues forming the glutathione binding site are conserved in the different classes, while those forming the substrate binding site vary considerably, leading to a wide substrate specificity (Johansson and Mannervik, 2001). Human GSTs consist of three families - cytosolic GSTs, mitochondrial GSTs and microsomal GSTs (Hayes *et al*., 2005). Cytosolic GSTs are further categorized into seven major classes - alpha (five members), mu (five members), pi (one member), theta (two members), zeta (one member), omega (two members), and sigma (one member). Members of the same class possess greater than 40 percent amino acid sequence identity and between classes, proteins have less than 25 percent sequence identity (Mannervik *et al*., 2005).

GSTs play an important role in cellular defense as they catalyze the conjugation of reduced glutathione with various electrophilic compounds such as the one present in tobacco smoke (Ishii *et al*., 1999). Associations between GST genotypes and disease phenotype may reflect a link between alleles and cytogenetic damage and specific mutations in target genes. GSTM1 and GSTT1 null cells are more susceptible to sister chromatid exchange following exposure to various electrophiles. Presumably genotypes, alone or in combination, should identify subjects who are detoxification deficient and consequently more likely to suffer formation of carcinogen DNA adducts and/or mutations (Ryberg *et al*., 1997). Thus, these two loci, GSTM1 mapped on chromosome 1p13.3 and GSTT1 on chromosome 22q11.23 are studied in particular because of their relevance in indicating susceptibility to cancer. Studies have shown that individuals who inherit the GSTM1 null genotype are not capable of conjugating and detoxifying specific substrate epoxide intermediates (Wiencke *et al*., 1990). Thus, the absence of the GSTM1 gene should increase cancer risk from environmental exposure while the presence of the intact GSTM1 gene would be protective for cytogenetic damage and carcinogen derived DNA adduct formation. GSTT1 has also

been involved in the glutathione-dependent detoxification. Similar to GSTM1, GSTT1 has significant activity towards epoxides, suggesting that individuals without both GSTM1 and GSTT1 may be at a particularly high risk of cancer (Wiencke *et al*., 1995).

In conclusion, in spite of numerous published data on breast cancer etiology, there is no proper consensus on these findings. This could be due to variation in the geographical, racial and ethnic distribution of the disease. There are also very few studies to confirm that the risk factors established in other parts of the world be the reason for incidence among the Mizo population. With rise in breast cancer incidence in this region, it is of great importance to identify the risk factors specific for this region.

CHAPTER 2

REVIEW OF LITERATURE

2. Review of Literature

2.1. Brief History of Breast Cancer

The first authentic accounts of breast cancer can be traced back to 3,000-2,500 B.C., in ancient Egypt (Breasted, 1930). In 400 B.C., Hippocrates described the progressive stages of breast cancer and associated the origin of breast cancer with cessation of menstruation (Ariel, 1987). In 200 A.D., Galen attributed the accumulation of black bile in the blood to cause breast cancer (De Moulin, 1983). In 1713, Bernardino Ramazzini in Italy, noted a higher frequency of breast cancer in nuns than in married women (Pope, 2004). In 1806, the Society for Investigating the Nature and Cure of Cancer published the findings of a questionnaire about the disease commenting, 'with regard to cancer, it is not only necessary to observe the effects of climate and local situation but to extend our views to different employments, as those in various metals and manufactures; in mines and collieries; in the army and navy; in those who lead sedentary or active lives; in the married or single; in the different sexes, and many other circumstances. Should it be proved that women are more subject to cancer than men, then we may enquire whether married women are more liable to have the uterus or breast affected; those who have suckled or those who did not; and the same observations may be made of the single' (Society for Investigating the Nature and Cure of Cancer, 1806). In 1842, Rigoni Stern in Italy, compared the incidence of cancer of the breast and uterus among married and unmarried females and showed the relationship of marital status to these cancers (Rigoni-Stern, 1842).

In 1915, a major contribution to breast cancer epidemiology came from an annual report of the Registrar-General of births, deaths and marriages in England and Wales. This report indicated that the mortality rate for breast cancer was markedly higher in single than in married women after the age of 45 (Stevenson, 1915). In 1926, Janet Elizabeth Lane-Claypon carried out the first modern case control study in United Kingdom comparing 508 breast cancer patients with 509 healthy women. This

study concluded that low fertility increases breast cancer risk (Lane-Claypon, 1926). In 1931, Lane-Claypon's study was replicated by JM Wainwright using a United States sample of 679 breast cancer cases and 567 unmatched controls. The 1926 United Kingdom study and 1931 United States study marked the beginning of a new era of etiologic research as it provided the first evidence from observational studies that parity, age at marriage, and artificial menopause were associated with breast cancer risk (Wainwright, 1931).

In 1866, Paul Broca wrote about the high prevalence of breast carcinoma in his wife's family and identified cases up to four generations from her family (Broca, 1866). This is the first of many reports that pointed out heritability of breast cancer and increased susceptibility for persons having positive family history of breast cancer (Claus *et al*., 1998). Key evidence that a single genetic mutation could cause heritable breast cancer risk came with the identification of a locus on chromosome 17q that was linked to disease susceptibility in specific families (Hall *et al*., 1990). In 1994, the BRCA1 gene was subsequently identified through positional cloning (Miki *et al*., 1994). During the same year, the second breast cancer susceptibility locus was localized to chromosome 13q12-13 by linkage studies of families with multiple cases of early-onset breast cancer that were not linked to BRCA1 (Wooster *et al*., 1994). The BRCA2 gene was cloned in 1995 (Wooster *et al*., 1995) and its complete coding sequence and exonic structure were described in 1996 (Tavtigian *et al*., 1996). ATM, BARD1, CHEK2, PALB2, TP53, PTEN, CDH1, STK11, RAD51C and RAD51D have also been recognized as breast cancer predisposition genes with a high to moderate risk. Most breast cancer predisposition genes participate in DNA damage repair pathways and cell cycle checkpoint mechanisms (Yoshimura *et al*., 2022).

The molecular mechanisms underlying the development of breast cancer are not completely understood. These tumours are likely to be caused by the interaction between many genetic and environmental factors. It was found that chemicals,

radiation and viruses inflict potential harm at certain regions on the hereditary material and their presence leads to impairment in the functionality of several genes (Cook *et al*., 1933; Riou *et al*., 1990). It was also observed that certain changes termed as germline mutations on this macromolecule were heritable. This further elicits the mechanistic approach regarding a better understanding of disease progression (Malik *et al*., 2009). With the advent of modern technologies, researchers and oncologists had a better understanding of cancer not only at the cellular but also at the macro and micro molecular levels. It is becoming more evident that in most cases, genotypic inheritance and lifestyle are probably inseparably intertwined. It is generally believed that the initiation of breast cancer is a consequence of cumulative genetic damages leading to genetic alterations resulting in activation of proto-oncogenes and inactivation of tumor suppressor genes, followed by uncontrolled cellular proliferation and/or aberrant programmed cell death or apoptosis. The role of reactive oxygen species (ROS) has been related to the etiology of cancer (Emerit, 1994).

2.2. Epidemiology

The statement that one out of every eight women will develop breast cancer in their lifetime indicates the gravity of this disease (Ferrini *et al*., 2015). Even though breast cancer account for the most number of all new cases of cancer in females, it ranks as the fifth most common cause of death, because of the relatively more favourable prognosis (mortality to incidence ratio, 0.35) making it the most prevalent cancer in the world today (Parkin, 2004). The global burden of breast cancer continues to rise with over one million new cases diagnosed and 400,000 deaths occurring each year in women. According to GLOBOCAN an online database that maintains cancer statistics for 185 countries, in 2020 there were 78 lakhs women alive who were diagnosed with breast cancer in the past 5 years; and 26 percent of all newly diagnosed cancer is cancer of the breast (Figure 6). In India, breast cancer accounted for 13.5 percent of all cancer cases and 10.6 percent of all deaths.

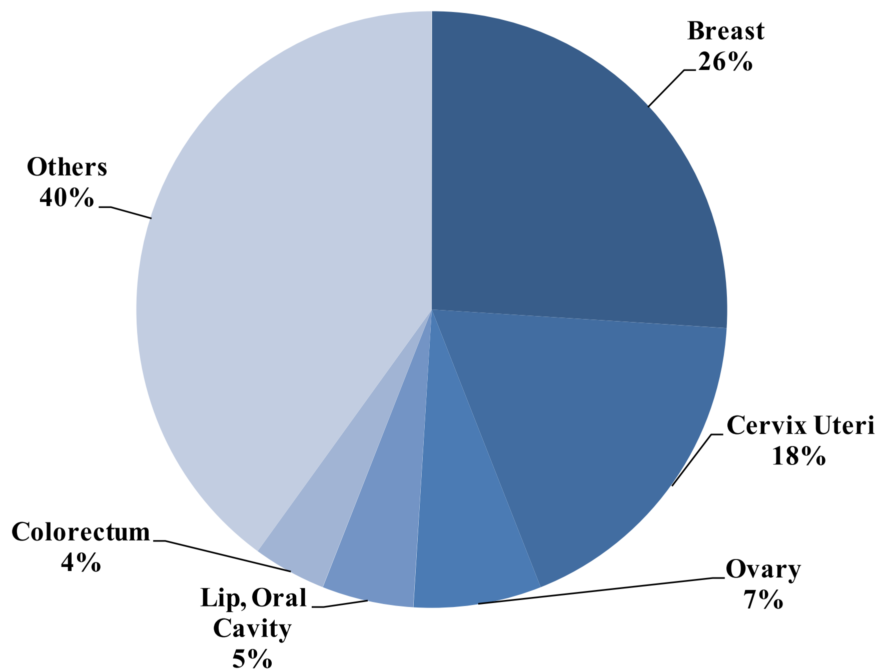


Figure 6. Percentage of new cases in females, Globocan 2020.

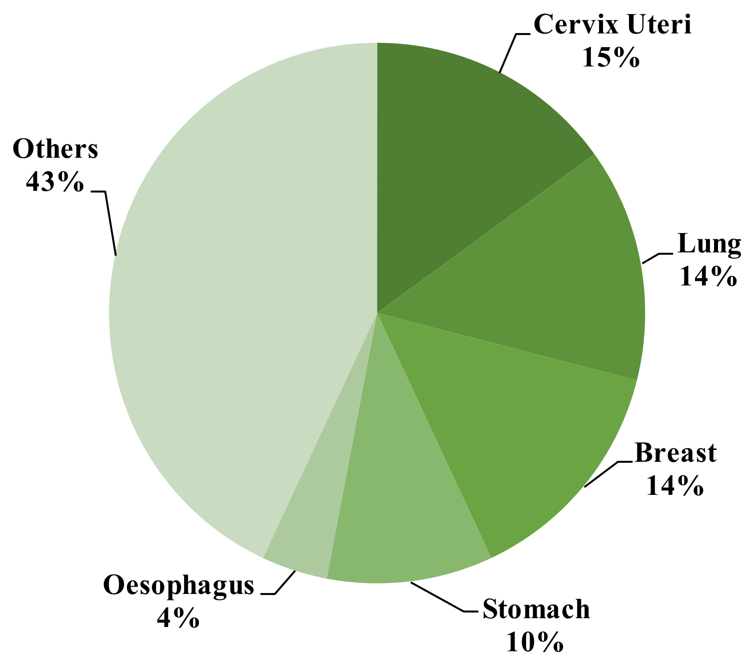


Figure 7. Percentage of new cases in females (Mizoram), National Cancer Registry Programme Report 2020.

According to PBCR 2012-2014 report, the highest age adjusted rate (AAR) of breast cancer were in Delhi (41.0), Chennai (37.9), Bangalore (34.4) and Thiruvananthapuram District (33.7). In Mizoram, breast cancer is at a crude rate (CR) of 15.8 and AAR of 19.9 per 100,000 population, with breast cancer being the third (13.5%) most common cancer among females after cervix uteri (15.4%) and lung (14.1%) as shown in Figure 7. In just Aizawl district, breast cancer is third (14.5%) most common after lung (17.8%) and cervix uteri (15.6%). In the other seven districts excluding Aizawl, breast cancer is the fourth (11.3%) most common after cervix uteri (16.3%), lung (13.3%) and stomach (12.1%). According to National Cancer Registry Programme (NCRP) 2020 report, the highest breast cancer incidence among all the North-Eastern PBCRs with AAR 30.7 is Aizawl district (Mathur *et al*., 2020).

Breast cancer incidence rate varies widely worldwide due to a range of socio-economic, reproductive, hormonal, nutritional and genetic factors which can be broadly classified into two factors, modifiable and non-modifiable (McPherson *et al*., 2000). Incidence is more closely associated to age than to any other risk factor, it increases rapidly during the reproductive years and then more slowly after about 50 years of age (Key *et al*., 2001). About 55 percent of the global burden is among developed countries, but incidence rates are rapidly rising in developing countries. While incidence rate is less than 40 per 100,000 women in most of the less developed countries, breast cancer is still the most common cancer among women in these countries (Parkin *et al*., 2005). A history of benign breast disease is also a well-established risk factor for breast cancer. Women with severe atypical epithelial hyperplasia have been found to have significantly higher risk compared to women who do not have any proliferative changes in their breast (McPherson *et al*., 2000). Numerous studies have been conducted to find the risk factors for breast cancer, some of the major recognized and suspected factors are listed as follows -

2.2.1. Reproductive Factors

The role of reproductive factors must be one of the most studied factors in breast cancer risk. Studies have revealed that a woman's breast undergoes many changes especially from puberty till menopause exposing it to high endogenous hormone levels (Key, 1999; Kabuto *et al*., 2000). Steroid hormones stimulate cellular replication and mitotic activity in breast epithelium which are believed to be crucial in the pathogenesis of mammary cancer. This high rate of cell division increases the frequency and likelihood of propagation of copying errors and DNA changes (Pike *et al*., 1983). Results from animal studies indicate that estrogen metabolites have genotoxic properties (Yager and Davidson, 2006). Lifetime exposure to endogenous sex hormones is determined by several variables including timing of menarche, age at first full term pregnancy, number of pregnancies, and age at menopause (Feigelson and Henderson, 1996). Cumulative lifetime exposure to estrogen is thought to be a key factor in determining breast cancer risk in women (Henderson *et al*., 1985).

2.2.1.1. Age at Menarche and Menopause

The milestone events that determine the period over which women are exposed to endogenous ovarian hormones have repeatedly been reported to influence breast cancer risk (Tulinius *et al*., 1978; Kvåle and Heuch, 1988; Hsieh *et al*., 1990). Late onsets of menarche as well as early menopause are associated with significant decrease in risk of 5 percent per year and 3 percent per year respectively (Collaborative Group on Hormonal Factors in Breast Cancer, 1997). The increased breast cancer risk associated with early age at menarche (< 12 years) (Kelsey *et al*., 1993; Bernstein *et al*., 1994; Berkey *et al*., 1999; Maurya and Brahmachari, 2022) is probably due to prolonged exposure of breast epithelium to estrogen with earlier onset of regular menstrual cycles and higher estrogen levels for longer years (Henderson *et al*., 1985; Apter *et al*., 1989). Urinary estrogens are significantly higher in girls who have early menarche than in those with normal menarche (Shi *et al*., 2010). Similarly, later age at menopause

maximises the number of ovulatory cycles and may therefore lead to increased risk. For every one year increase in the age at menopause, the risk of breast cancer increases by approximately 3 percent (Collaborative Group on Hormonal Factors in Breast Cancer, 1996). One of the most compelling pieces of evidence regarding the influence of endogenous hormones on breast cancer risk is found in the levelling off in the age-specific incidence curve of breast cancer after menopause when ovarian production of steroid hormones ceases. Breast cancer risk in women with natural menopause before the age of 45 years is half compared with women who stop menstruating after the age of 55 years (Pike *et al*., 2004). There is also convincing evidence of an age-dependent protective effect of early menopause surgically induced by bilateral oophorectomy (Brinton *et al*., 1988; Li *et al*., 2016).

2.2.1.2. Childbearing

The exact mechanism by which an early first birth protects against breast cancer remains incompletely understood, but has primarily been attributed to shortening of the time window of high susceptibility to carcinogenic transformations (Russo *et al*., 2000). A woman's risk of breast cancer appears to be related to timing of first birth and age at subsequent births. Higher parity and early age at first birth have both been associated with decreased lifetime incidence of breast cancer (Kelsey *et al*., 1993; Chie *et al*., 2000; Pathak *et al*., 2000). Studies in India among Parsi women, who on average are wealthy, marry late, and have few children, have an age-adjusted incidence rate of breast cancer that is more than twice that of Hindu women living in the same geographical area, who as a group are poorer, marries earlier, and have more children (Jussawalla *et al*., 1981).

In most Western countries, there is a social gradient in breast cancer risk with markedly higher incidence in women with high education compared to women with low education (Faggiano *et al*., 1997; Shack *et al*., 2008). For a woman to have higher education, childbearing is usually delayed in most cases. Breast cancer risk is half in

women having their first child before the age of 20 compared to those having their first child after the age of 30 (Russo *et al*., 2000). There is also some evidence that the interval between age at menarche and age at first birth may be relevant to breast cancer risk (Andrieu *et al*., 2000; Clavel-Chapelon and E3N Group, 2002; Li *et al*., 2008). One study reported that the length of this interval was related to risk, particularly to risk of hormone receptor-positive tumors (Li *et al*., 2008).

A study in 1970 observed protective effect of parity attributing to an earlier age at first childbirth in women with many children (MacMahon *et al*., 1970). It is now estimated that for each additional year of age at first birth, the risk of premenopausal breast cancer increases by 5 percent, and increases by 3 percent for breast cancers diagnosed after menopause (Clavel-Chapelon and Gerber, 2002). Compared to nulliparous women, women with a first full-term pregnancy before age 20 years have about half the risk of breast cancer (Kelsey *et al*., 1993). Women with an older age at first birth (≥35 years) have the same risk of breast cancer as nulliparous women. In a reanalysis of MacMahon's data, older age at any birth was found to be an independent risk indicator (Trichopoulos *et al*., 1983). In an Italian case control study, breast cancer risk increased 0.7 percent per year when subsequent births were delayed (Decarli *et al*., 1996). Even though association and patterns of childbirth cannot exclude genetic and lifestyle influences, each pregnancy and the timing of birth act as markers for cumulative exposure to ovarian hormones and possibly other risk modifying factors, some yet to be identified.

2.2.1.3. Parity

It has long been recognized that parity reduces the risk of breast cancer (Kelsey *et al*., 1993). In 1713, Ramazzini of Padua observed what appeared to be an epidemic of breast cancer among nuns (Pope, 2004). One hundred years later, it was noted that breast cancer was at least three times as frequent in nuns as in other women (Rigoni-Stern, 1842). In a study published in 1926, Lane-Claypon reported an association

between reproductive history and breast cancer risk (Lane-Claypon, 1926), findings which were confirmed in the early 1930s (Wainright, 1931), and later in British vital statistics data which revealed a high breast cancer mortality in unmarried and childless women (Gilliam, 1951). The risk reduction appears to be greatest among women with high parity, where the risk reduction due to breastfeeding may be as great as 50 percent (Romieu *et al*., 1996) and among premenopausal women with lactation durations ≥2 years, where the breast cancer risk reduction may be 30 percent (Newcomb *et al*., 1994). There is also evidence that the timing of pregnancy is relevant to breast cancer risk. Compared to single women, the risk of breast cancer is lower in older married women, but not in younger married women, with an approximate crossover of the effect around age 40 (Janerich and Hoff, 1982; Pathak *et al*., 1986). There are other reports of a higher breast cancer risk among young parous compared to young nulliparous women (Woods *et al*., 1980; Layde *et al*., 1989) and an increased risk of breast cancer in the years following childbirth (Bruzzi *et al*., 1988; Layde *et al*., 1989; Williams *et al*., 1990). Pregnancy induces both transient and permanent structural changes in the breast tissue of laboratory animals (Russo *et al*., 1982, 1990). It appears that the effect of parity is determined by the age of occurrence of component pregnancies and that the closer the births are together, the lower the risk. A likely explanation is that pregnancies occurring close together in time provide less time for breast cells to accumulate DNA damage and that every new pregnancy affords additional protection by recruiting more of the remaining undifferentiated cells (Russo and Russo, 1993).

There is much debate whether an incomplete pregnancy affects future breast cancer risk. Based on findings from animal studies, it has been hypothesized that an increase in breast cancer risk may follow if the hormonal surge occurring during the first trimester is not followed by the protective components of breast tissue maturation and terminal differentiation of lobular structures during the second and third trimester (Russo *et al*., 1982). Findings from case control studies indicated that

induced abortions were associated with an increased risk of breast cancer (Michels and Willett, 1996). However, collective evidence to date points to no association between pregnancy interruption and subsequent breast cancer risk. In 2003, a National Cancer Institute expert panel concluded that neither spontaneous nor induced abortions are associated with an increased risk of breast cancer (http:// www.cancer.gov/cancerinfo/ere-workshop-report).

2.2.1.4. Breastfeeding

In the 1920s, it was observed that the children of women with breast cancer were less likely to have been breastfed for 1 year than the children of control women (Lane-Claypon, 1926). Lactation decreases risk of breast cancer in parous women, although the overall reduction in risk varies substantially within the population studied (Lipworth *et al*., 2000). The relative risk for breast cancer decreases by 4.3 percent for every 12 months of breastfeeding (Collaborative Group on Hormonal Factors in Breast Cancer, 2002) and decreased risk with prolonged lactation (Lipworth *et al*., 2000). The risk reduction appears to be greatest among women with high parity, where the risk reduction due to breastfeeding may be as great as 50 percent (Romieu *et al*., 1996) and among premenopausal women with lactation durations ≥2 years, where the breast cancer risk reduction may be 30 percent (Newcomb *et al*., 1994). The magnitude of the decline was consistent across age at breast cancer diagnosis, race or ethnicity, different reproductive patterns, and various personal characteristics. This led to the conclusion that the limited time women in developed countries breastfeed are likely the reasons for high breast cancer incidence (Kelsey *et al*., 1993). Breastfeeding is hypothesized to reduce the risk of breast cancer through differentiation of breast tissue and reduction of the lifetime number of ovulatory cycles. Breastfeeding may result in further terminal differentiation of the breast epithelium, making it more resistant to carcinogenic change. However, some reviews have consistently found no association that breastfeeding reduces risk. There is no solid consensus about the relationship between breastfeeding

and breast cancer. Expanded consideration of possible confounders for this relationship is required to determine if breastfeeding is protective and how protection might be conferred (Yang and Jacobsen, 2008).

2.2.2. Family History and Medical History

Hereditability of cancer can be viewed from several angles. Inherited cancer may refer to all cancer cases bearing established causal genetic mutations, and it may refer to cases in families with multiple cancer cases, although no common causal genetic trait has been identified. Since 1860, breast cancer has been recognized to cluster in families as described by the French surgeon Paul Broca (Lynch *et al*., 2008). BRCA1, BRCA2, ATM and TP53 are the most common high penetrance genes exhibiting allelic variants predisposing to hereditary breast cancer (Easton *et al*., 1993). Germline mutations in BRCA1 and BRCA2 genes are the main genetic and inherited factors for breast cancer. It is estimated that about five to ten percent of all breast cancer cases are caused by mutations in these high-risk genes (Easton, 2002; Lux *et al*., 2006; Lynch *et al*., 2008). These mutations are important in developing early onset and increasing the risk of familial breast cancer, and responsible for 90 percent of hereditary cases (Mahdavi *et al*., 2019). In many families no such pattern can be found, but the history is still indicative of a kind of genetic predisposition. But this susceptibility explains only a small fraction of the familial risk and a much smaller fraction of 5-10 percent of breast cancer cases and decreases markedly with age; approximately 33 percent of cases age 20-29 years compared with approximately 2 percent of cases age 70-79 years (Claus *et al*., 1996). Several studies have been undertaken to find other high penetrance breast cancer susceptibility genes. Genetic variants in CHEK2, PTEN, STK11, CDH1, NBS1, RAD51C, RAD51D, BARD1, PALB2, XRCC2 and XRCC3 have also been implicated in breast cancer risk. Some of these genes are involved in multiple cancer syndromes like Li-Fraumeni (TP53), Peutz-Jeghers (STK11/LKB1) and Cowden syndrome (PTEN) (Ko and Prives, 1996; Nelen *et al*., 1996; Hemminki

et al., 1998; Sigal and Rotter, 2000; Gasco *et al*., 2003). In contrast to familial cases, low-penetrance genes contribute to sporadic cases of breast cancer that usually appear unilaterally and have a relatively late age at diagnosis (Rebbeck, 1999).

Some studies suggest that hypertension increases the risk of all malignancies (Grossman *et al*., 2001) via hypothesized pathways relating to abnormalities of vascular smooth muscle proliferation, carcinogen binding to DNA, or angiogenesis (Felmeden and Lip, 2001). While some early studies reported a positive association between hypertension and breast cancer (Largent *et al*., 2006), some studies showed no differences even when a follow-up of up to 27 years was done (Peeters *et al*., 2000; Manjer *et al*., 2001; Lindgren *et al*., 2007). A meta-analysis of 30 studies, with totally 11643 cases showed statistically significant association between hypertension and increased breast cancer risk. A study found increased risk in postmenopausal women than in premenopausal women and Asian population (Han *et al*., 2017). Treatment for hypertension has also been associated with breast cancer risk in several studies although the evidence is inconsistent (Grossman *et al*., 2001). Most observational epidemiologic data do not support an association between antihypertensive use and breast cancer. Furthermore, several large randomized clinical trials showed no association between risk of any cancer and antihypertensive treatment (Lindholm *et al*., 2001; ALLHAT Collaborative Research Group, 2002).

History of diabetes, a condition marked by sustained, high insulin levels, has been associated with increased breast cancer risk (Wu *et al*., 2007). The pathways by which diabetes might cause breast cancer involves the insulin pathway, activation of the insulin-like growth factor pathway, and altered regulation of endogenous sex hormones (Wolf *et al*., 2005). The latter two pathways are thought to be key mechanisms linking obesity and breast cancer. In one study wherein type 1 and 2 diabetes was distinguished, the risk of breast cancer was increased by statistically significant 17 percent among postmenopausal women with type 2 diabetes (Michels *et al*., 2003).

Although this is the largest study to date with the longest follow-up (over 22 years), self-reported body weight and lack of data on central obesity could have resulted in residual confounding. In the Cancer Prevention Study II which had a follow-up of 16 years, the relative risk for breast cancer mortality among women with diabetes was a statistically significant with a relative risk of 1.27 (Coughlin *et al*., 2004). A meta-analysis of 40 studies showed there was significant increase in risk associated with diabetes in women. Even after sub-grouping by type of diabetes, the association was unchanged with type 2 but nullified with gestational diabetes (Hardefeldt *et al*., 2012). Data from some case control and cohort studies suggest that diabetes carries a moderate increase in the risk of breast cancer (Xue and Michels, 2007).

2.2.3. Lifestyle Factors

Lifestyle factors are modifiable risk factors that can have an important role in primary breast cancer prevention, breast cancer treatment, and tertiary breast cancer prevention. It is estimated that 90-95 percent of breast cancer cases are connected to environmental factors and lifestyle (Castelló *et al*., 2015). For over three decades, World Cancer Research Fund and American Institute for Cancer Research (WCRF/AICR) has been at the forefront of synthesizing, interpreting, and evaluating the accumulated evidence on the relationship of diet, nutrition, physical activity, and weight with cancer risk (Wiseman, 2008). In the United States alone, approximately 40 percent of all cancer cases could be prevented through health-related choices such as vaccinations and modifiable lifestyle factors, including body weight, physical activity level, alcohol intake, diet, sun exposure, and tobacco use (McCullough *et al*., 2011; Islami *et al*., 2018). Some of the major lifestyle factors that are said to increase breast cancer risk are -

2.2.3.1. Anthropometric Factors

Measurements of a human body include height, weight, body mass index (BMI)

and other proportions of the body. Even though there is as yet no clear explanation for the connection between height and breast cancer risk, it has been suggested that within populations a 10 cm increase in height corresponds to a 10 percent increase in risk (Hunter and Willett, 1993). Since estrogen plays a key role both in breast cancer development and human growth regulation, the growth spurt has been modestly suggested to influence risk. Estrogen stimulates the pubertal growth spurt and mutations in the ESR1 gene (coding for ERα) have been reported to delay fusion of the epiphyseal plates at puberty (Simm *et al*., 2008; Emons *et al*., 2010). An association between body height and mutations in ESR1 has also been found (Dahlgren *et al*., 2008), and might point towards a more hormone related link.

Body mass index (BMI) and waist-to-hip ratio (WHR) are the most popular measurement for body fat. Several studies have found obesity to be associated with an increased breast cancer risk among postmenopausal women, but in premenopausal women either unrelated or related to a reduced risk (Lubin *et al*., 1985; Willett *et al*., 1985; Chu *et al*., 1991; Pathak and Whittemore, 1992; Ballard-Barbash, 1994; Franceschi *et al*., 1996; Huang *et al*., 1997; Sonnenschein *et al*., 1999). High BMI, an indicator of obesity, has been suspected to increase risk in postmenopausal women (Hunter and Willett, 1993; Reeves *et al*., 2007). Obesity has been related to both higher endogenous estrogen levels and circulating estrogen in the adipose tissue (Hunter and Willett, 1993; Key *et al*., 2003). In premenopausal women, this connection is unclear (Renehan *et al*., 2008). However, some studies found BMI is inversely associated with risk among premenopausal women (McTiernan, 2003) as obese premenopausal women are more likely to have irregular menstrual cycles and ovulatory infertility (Rich-Edwards *et al*., 1994). Study on association of waist-to-hip ratio with breast cancer risk show significant results (Ballard-Barbash, 1994; Mannisto *et al*., 1996; Hall *et al*., 2000).

2.2.3.2. Physical Activity

Few established breast cancer risk factors are modifiable. However, increasing physical activity and maintaining weight during a woman's adult years offer both individual and population-based opportunities for lowering women's risk of breast cancer (Bernstein *et al*., 1994; De Cree *et al*., 1997). Consistent evidence from epidemiologic studies links physical activity after diagnosis with better breast cancer outcomes as well (Ibrahim and Al-Homaidh, 2011; Chlebowski, 2013). The relationship between physical activity, anthropometric factors, and breast cancer risk may be mediated by several pathways including the steroid hormone, insulin, and insulin-like growth factor pathways. The link between estradiol and breast cancer has been supported by in vitro (McManus and Welsch, 1984; Laidlaw *et al*., 1995) and in vivo (Chang *et al*., 1995) studies showing that estradiol increases the mitotic activity of breast epithelial cells. Physical activity appears to have a direct physiological effect on steroid hormone levels, most clearly during the pubertal and premenopausal stage. Increased physical activity has been directly associated with reduced circulating levels of endogenous estradiol and progesterone among normally cycling women (Shangold *et al*., 1979; Ellison and Lager, 1986). It has an indirect effect on exposure to ovarian steroid hormones, in that high level of moderate and vigorous physical activity result in delayed menarche, irregular or anovulatory menstrual cycles, a shortened luteal phase, and in the extreme, secondary amenorrhea (Warren, 1980; Frisch *et al*., 1981; Bernstein *et al*., 1987).

Studies on physical activity and circulating hormone levels in postmenopausal women have not given consistent results (Newcomb *et al*., 1995; Verkasalo *et al*., 2001; Atkinson *et al*., 2004). Breast cancer risk is decreased most with recreational and household activities after menopause. Multiple mechanisms could explain associations between postmenopausal breast cancer, estrogen levels, and physical activity. The first relates to BMI, increase of which is related to breast cancer risk (World Cancer

Research Fund and the American Institute for Cancer Research, 2007). This relation might exist in part because after menopause, ovarian estrogen production ceases and adipose tissue becomes a key endogenous source of circulating estrogen (Kendall *et al*., 2007). Hence, by reducing body fat through exercise, estrogen levels may decrease resulting in a lower risk of breast cancer. Levels of adipokines that influence estrogen biosynthesis can also be altered with weight loss (Cleary and Grossmann, 2009).

2.2.3.3. Sleep Cycle and Circadian Disruption

Several studies have investigated a potential link between night shift work and the development of breast cancer (Davis *et al*., 2001; Schernhammer *et al*., 2001, 2006) and have shown an increased risk among women who work in occupations that typically involve some degree of shift work (Megdal *et al*., 2005). The release of nearly all hormones exhibits a circadian timing patterned on approximately a 24-hour cycle. Agents that disrupt circadian rhythm may also alter endocrine function and thereby the regulation of reproductive hormones. Sleep exerts a profound effect on endocrine function and hormones such as melatonin and cortisol (Czeisler and Klerman, 1999). Clinical studies in women with breast cancer showed that they had much less night-time melatonin levels in urine than a control group of women with no breast cancer (Schernhammer and Hankinson, 2005). It was reported that melatonin is reduced and estrogen elevated in nurses with a history of rotating night shifts (Schernhammer *et al*., 2004). The Finnish Twin Cohort Studies concluded that the risk of breast cancer was lower in women who sleep longer (>9 hr) compared to average sleepers (7-8 hr) (Verkasalo *et al*., 2005).

2.2.3.4. Alcohol

There is substantial evidence to support the association of increased breast cancer risk with alcohol consumption. In a pooled analysis of six cohort studies, alcohol was shown to increase breast cancer risk linearly with alcohol consumption from 1-5

drinks/day (Smith-Warner *et al*., 1998). However, only a modest 15 percent increase in risk was seen in a study of alcoholic women (Kuper *et al*., 2000). A collaborative reanalysis of data from 53 epidemiological studies (Hamajima *et al*., 2002) estimated that the relative risk of breast cancer increased by 7.1 percent for each additional 10 g per day intake of alcohol, i.e., for each extra unit or drink of alcohol consumed daily. The association was seen in both premenopausal and postmenopausal women but does not vary by type of alcoholic beverage (Smith-Warner *et al*., 1998; Tjonneland *et al*., 2007), and does not seem to depend on drinking frequency (Tjonneland *et al*., 2003; Horn-Ross *et al*., 2004). Recent alcohol intake seems to be more relevant than past intake. Alcohol intake in adolescence is not associated with subsequent breast cancer risk (Holmberg *et al*., 1995, Marcus *et al*., 2000). A controlled feeding trials have shown that moderate alcohol intake increases circulating estrogen levels in both premenopausal (Reichman *et al*., 1993) and postmenopausal (Dorgan *et al*., 2001) women.

Recent advances have indicated that alcohol consumption is strongly related to estrogen receptor positive (ER+) breast cancers (Deandrea *et al*., 2008). Alcohol-associated impact on breast cancer appears to be effective in ER+ invasive lobular carcinoma, but not in ER+ invasive ductal carcinoma (Li *et al*., 2010). Studies suggest that for alcohol drinkers, interactions with GSTM1 and GSTT1 deletion polymorphisms may play an important role in individual susceptibility to breast cancer (Helzlsouser *et al*., 1998; Park *et al*., 2000). In vitro, addition of alcohol to breast cancer cells resulted in ER signalling and cell proliferation of ER+ but not ER- cells (Fan *et al*., 2000; Singletary *et al*., 2001).

The mechanism underlying the carcinogenic effect associated with alcohol is not completely understood. However, an increase in the estrogen level in women consuming alcohol has been hypothesized (Reichman *et al*., 1993). Some proposed mechanisms include alcohol induced production of ROS (Wright *et al*., 1999),

and increased adduct formation, possibly due to decreased protein expression of detoxification enzymes (Barnes *et al*., 2000). It is also believed that alcohol intake increases mammary tissue exposure to estrogen, induces mutagenesis through its metabolites, increases free radical-mediated DNA damage, and may influence DNA metabolism and gene expression by affecting one-carbon metabolism (Dumitrescu and Shields, 2005; Seitz and Stickel, 2007).

2.2.3.5. Tobacco

Tobacco is known to contain a variety of compounds that are carcinogenic. Cigarette smoking is the most common form of tobacco use worldwide. Tobacco consumption has been clearly implicated in the causation of many cancer types. However, studies on breast cancer risk have reported positive, inverse and null associations. Despite mixed result, there is growing evidence that smoking may slightly increase the risk of breast cancer. In a meta-analysis, current and former smoking were weakly associated with breast cancer risk; a stronger association was observed in women who initiated smoking before first birth (Gaudet *et al*., 2013). Statistically significant effects have been seen for early age at starting, and for heavy, current, and passive smoking. In some studies, there is significant increase in breast cancer risk in never smokers with longterm exposure to passive smoking, while other scientist rejects the evidence of this association as inconsistent (Ambrosone and Shields, 1999; Wartenberg *et al*., 2000). In a study where non-smoking women exposed to passive or secondhand smoke were excluded, there was evidence of positive associations between breast cancer and cigarette smoking (Morabia, 2002). In addition, studies suggest that the risk of breast cancer associated with smoking might be increased for premenopausal women (Khuder *et al*., 2001) or women who started smoking in their mid-teens or earlier, or before their first full-term pregnancy (Khuder *et al*., 2001; Terry and Rohan, 2002; Cui *et al*., 2006; Ha *et al*., 2007). Similarly, women who inherited specific variants in genes involved in the metabolism of carcinogens found

in tobacco might experience higher risks associated with smoking cigarettes (Morabia, 2002; Terry and Rohan, 2002; Slattery *et al*., 2008).

Tobacco smoke consists of more than 7000 chemical compounds, and more than 60 of these are known carcinogens (Das, 2003). These are transported by the blood stream, deposited and metabolically activated in the breast and surrounding adipose tissues to the breast, which can further be detected in the nipple discharge or as smoking specific DNA adducts in breast tissue (Terry and Rohan, 2002). The most important carcinogens in tobacco smoke are polycyclic aromatic hydrocarbons (PAHs), aryl amines, heterocyclic aromatic amines (HAAs), and N-nitrosamines (Ambrosone and Shields, 1999; Bartsch *et al*., 2000). The ingested or inhaled PAHs are converted to water soluble derivatives mainly via oxidative activation by cytochrome P450 1A1 (CYP1A1) followed by detoxification by phase II enzymes such as glutathione s-transferases (GSTs). PAHs have been shown to be mutagenic to breast cell lines, and as lipophilic compounds they are stored in adipose tissues (Li *et al*., 1996).

Smokeless tobacco products contain more than 3000 chemicals and 28 numerous carcinogens. DNA binding and mutations are among the mechanisms clearly implicated in carcinogenesis due to smokeless tobacco use. Smokeless tobacco generally comes in the form of chewing tobacco. Smokeless tobacco also generates reactive oxygen species, oxidative stress,and associated DNA fragmentation in laboratory experiments. The major and most abundant group of carcinogens are the non-volatile alkaloid-derived tobacco-specific N-nitrosamines (TSNA) and N-nitrosoamino acids. Other carcinogens reportedly present in smokeless tobacco include volatile N-nitrosamines, certain volatile aldehydes, traces of some polynuclear aromatic hydrocarbons such as benzo[a]pyrene, certain lactones, urethane, metals, polonium-210 and uranium-235 and -238 (Brunnemann and Hoffmann, 1992). Compared to smoking, higher levels of nicotine can enter systemic circulation from smokeless tobacco which indicate a much more potent effect through this route (Li *et al*., 2018). *Tuibur*, a unique form of aqueous

extract of tobacco exhibited significant toxicity by reducing the root growth of Allium bulbs and inducing tumor formation in the roots. This form of smokeless tobacco is most commonly used in the north eastern part of India like Manipur, Mizoram, Sikkim and Tripura. It is used for gargling, or sipped and spitted out only after it becomes diluted (Mahanta *et al*., 1998).

2.2.3.6. Betel Quid

Betel quid can contain a variety of ingredients and combinations depending on different parts of the world. Usually, it contains a mixture of areca nut (*Areca catechu*), catechu (*Acacia catechu*) and slaked lime (calcium oxide and calcium hydroxide) and several condiments according to taste preference, wrapped in betel leaf (*Piper betle*), some add tobacco (Nair *et al*., 2004). In Mizoram, the most common preparation used in betel quid is a mixture of areca nut wrapped in betel leaf with a paste of slaked lime. In vitro and in vivo experiments have shown that betel quid consumption can cause micronuclei and DNA adducts formation, chromosomal aberrations, allelic imbalances and sister chromatid exchange in oral mucosa cells (IARC, 2004). Calcium hydroxide a major content of slaked lime in the presence of areca nut is responsible for the formation of ROS (reactive oxygen species) known to cause oxidative damage in the DNA of buccal mucosa cells of betel quid chewers (Nair *et al*., 2004). A study in Assam, India, reveals that betel quid chewers have higher risk of having breast cancer than the non-chewers (Rajbongshi *et al*., 2015). In a case control study among the Mizo population betel quid chewing was found to be a significant risk factor for developing breast cancer. Multifactor dimensionality reduction identified betel quid chewing as the single main risk factor and women with betel quid chewing history have five times the risk of developing breast cancer (Kaushal *et al*., 2010). Higher total number of genomic alterations were seen in breast cancer tumors of betel quid than to non-chewers (Kaushal *et al*., 2012).

2.2.4. Diet and Nutrition

The human diet contains a great variety of natural carcinogens and anticarcinogens (Sugimura, 2000). Many of these may act through the generation of oxygen radicals, which in turn may lead to DNA damage. There is an almost universal agreement that diet or nutritional practices in some form must play a role in establishing breast cancer risk. However, no specific component of the adult diet and no particular nutrient have been consistently associated with breast cancer risk (World Cancer Research Fund and American Institute for Cancer Research, 2010). The results of a large meta-analysis of 26 published studies from 1982 to 1997 (Gandini *et al.*, 2000) and of a pooled analysis of 8 cohort studies (Smith-Warner *et al.*, 2001b) suggest that fruit and vegetable consumption during adulthood is not significantly associated with reduced breast cancer risk. A pooled analysis of individual data from seven prospective studies in four countries comprising 337,819 women and 4980 breast cancers also suggested a lack of association between total fat, saturated fat, mono and poly unsaturated fat intake and breast cancer risk (Hunter *et al.*, 1996).

However, some studies have shown that a high intake of fat, especially that of poly unsaturated fatty acids, has been shown to increase breast cancer risk (Bartsch *et al.*, 1999), while intake of fruits and vegetables, sources of natural antioxidants, has been shown to decrease the risk (Lee, 1999; McKeown, 1999). Consumption of meat has been associated with increased breast cancer risk in some, but not all studies (Zheng *et al.*, 1998). Dietary fat has long been suspected to be the reason for this association, but recent studies support the role of heterocyclic aromatic amine (HAAs) found in well-done meat. HAAs require metabolic activation by N-acetyltransferase to be able to exert their harmful effects (Hein *et al.*, 2000; Hein, 2000). Soy, or more specifically genistein, with a chemical structure like steroidal estrogens, has been shown to have both anti-carcinogenic and cancer promoting effects (Bouker and Hilakivi-Clarke, 2000).

The Mizos are a unique group of people differing from mainland India in their culture, lifestyle and dietary habits. Some of the indigenous foods of the Mizos contained ash filtrate (*ching-al*), smoked or sun-dried or fermented meats and vegetables. Studies on the association of stomach cancer risk with these traditional food reveals that frequent consumption of fermented pork fat (*sa-um*), smoked dried salted meat and fish elevates risk. The use of soda or ash filtrate, used as a food additive, increased the risk of stomach cancer (Phukan *et al*., 2006). Studies reveals that smoked meat, smoked fish and soda also increased lung cancer risk (Phukan *et al*., 2014). Even though there is no published data on association of breast cancer risk among the Mizos with fermented soyabean (*bekang*), a study in a Japanese population-based cohort showed no association with risk (Shirabe *et al*., 2021).

2.2.5. Exogenous Hormones - Oral Contraceptives and Postmenopausal Hormone

Sex hormones have become one of the most widely used drugs among women. It is not possible to study the effects of single hormone as many are used in the same patient either in combination or consecutively. Therefore, risks are generally assessed in relation to the therapeutic goal of the treatment, i.e., oral contraception or hormone replacement therapy (HRT). Results from studies on the role of oral contraceptives to breast cancer proneness have been somewhat conflicting. However, data from 54 studies concluded that current use of oral contraceptives poses a slight (24%) increase in the risk, which disappears 10 years after the cessation of use (Collaborative Group on Hormonal Factors in Breast Cancer, 1996). The importance of progesterone in breast cancer risk has been highlighted by several recent observational studies and most notably the Women's Health Initiative randomized trials that have shown that combined estrogen plus progestin hormone therapy increases breast cancer risk while use of estrogen alone does not (Ross *et al*., 2000; Chlebowski *et al*., 2003).

Results from 51 studies indicated that risk of having breast cancer is slightly increased in women using HRT (Collaborative Group on Hormonal Factors in Breast Cancer, 1997). A 35 percent increase in risk was seen for women who have used HRT for 5 years or longer, being comparable with the effect of delaying menopause, and the increase largely disappeared 5 years after terminating the use of hormones. The combined estrogen-progestin regimen is associated with greater increase in risk than estrogen alone (Schairer *et al*., 2000). Moreover, in some studies, cancers in women who have ever used HRT tend to be less advanced at diagnosis and biologically less aggressive than those in never users (Holli *et al*., 1997), but contrasting results exist (Stallard *et al*., 2000). The overall mortality among HRT users has been shown to be lower but the benefit diminishes with longer duration of use (Grodstein *et al*., 1997).

Xenoestrogens include pesticides, dyes, pollutants, plasticizers and food preservatives that have estrogen-like effects, and they have been shown to have a role in the etiology of breast cancer (Davis *et al*., 1993; Dees *et al*., 1997b; Safe, 1997; Garner *et al*., 2000a; 2000b; Spink et al., 2000). Xenoestrogens have also been called endocrine disrupters because they interfere with the actions of endogenous estrogen. For instance, catechol metabolites of polychlorinated biphenyls (PCBs) have been suggested to alter estrogen metabolism by inhibiting the inactivation of carcinogenic estrogen metabolites (Garner *et al*., 2000a).

2.2.6 Environmental and Occupational Exposures

It has been widely suggested that environment may play an important role in increasing breast cancer incidence. Environmental pollutants such as organochlorines, polycyclic aromatic hydrocarbons, dioxins, and bisphenol A and extremely low frequency (ELF) magnetic fields have been linked with breast cancer risk in animal studies and may plausibly be associated with risk in humans. Exposure to some naturally occurring trace elements and heavy metals are suspected to influence breast cancer risk.

2.2.6.1. Environmental Pollutants

The International Agency for Research on Cancer (IARC) considers several environmentally abundant chemicals, chemical compounds, and their metabolites to be either known (IARC, 1997b) or suspected (IARC, 1997a) human carcinogens. Among these, organochlorines, polycyclic aromatic hydrocarbons (PAHs), dioxins, and bisphenol A (BPA) have received particular attention with respect to breast cancer. The specific mechanisms by which exposure to environmental pollutants could impact breast cancer risk are varied but their persistence in the environment and their tendency to accumulate in adipose tissue, including the fatty tissue in the breast is common. Concerns that exposure to these pollutants could influence risk stem primarily from the fact that many of these chemicals are 'endocrine disruptors,' mimicking or blocking the effects of specific hormones (Rudel *et al*., 2007). Since some of these pollutants mimic the activity of estrogen, it is hypothesized that they could influence the initiation or progression of breast cancer in humans through estrogenic effects (Soto *et al*., 1995; Connor *et al*., 1997; Shekhar *et al*., 1997).

Dichloro-diphenyl-trichloroethane (DDT) and its associated compounds have been classified as Group 2B carcinogens (possibly carcinogenic) by IARC (IARC, 1997a). It is ubiquitous in nature and accumulates in the food chain, particularly in fish and fatty foods. It has estrogenic effects (Soto *et al*., 1995; Dees *et al*., 1997a; Shekhar *et al*., 1997); studies have documented that DDT can regulate estrogen receptor (ER)-mediated cellular responses and stimulate cell cycle progression in ER-positive (ER+) breast cancer cell lines (Dees *et al*., 1997a; Shekhar *et al*., 1997). Polycyclic aromatic hydrocarbons (PAHs) are formed as by-products of combustion from tobacco smoke, air pollution, vehicle exhaust, and smoked or grilled meat and fish (Brody *et al*., 2007). Some PAHs are considered environmental estrogen, although their estrogenic effects are generally weak (Santodonato, 1997).

Although some studies provide no strong evidence of gene-environment interactions, there is some evidence to suggest that women with a GSTM1 null phenotype (Rundle *et al*., 2000) may be susceptible to the effects of PAH exposure. Studies have suggested that exposure to high levels of PAHs in early childhood (Bonner *et al*., 2005), at the time of first birth (Nie *et al*., 2007), or in the past 10-20 years (Lewis-Michl *et al*., 1996) could increase risk of postmenopausal breast cancer; associations with risk of premenopausal breast cancer are somewhat less consistent (Bonner *et al*., 2005; Nie *et al*., 2007). Bisphenol A (BPA) is an important monomer in the production of the epoxy resins that line food and beverage cans and in the production of the shatter proof polycarbonate plastics that are used in a wide variety of household products and devices. Studies in mouse and rat models have suggested that in utero exposure to BPA results in alterations in the architecture of the adolescent and adult breast. Mammary glands in animals prenatally treated with BPA have an increased number of undifferentiated epithelial structures, more progesterone receptor-positive (PR+) epithelial cells, decreased apoptosis and enhanced sensitivity to estradiol (Durando *et al*., 2007; Murray *et al*., 2007; Moral *et al*., 2008).

2.2.6.2. Radiation – Ionizing and Non-Ionizing

It is clearly established that exposure to ionizing radiation is an important risk factor for breast cancer. The most important information available regarding association of ionizing radiation with breast cancer risk comes from studies of the survivors of the atomic bombings of Hiroshima and Nagasaki (Land *et al*., 2003) and long-term follow-up of cohorts of people receiving radiation exposure for medical reasons (Horwich and Swerdlow, 2004). Age at exposure appears to be an important risk determinant and exposure around the time of puberty conferring the highest risk (Goss and Sierra, 1998). Ionising radiation has been shown to increase breast cancer risk among female flight attendants, nurses, chemists and insulators (John and Kelsey, 1993; McCormick, 1999; Weiderpass *et al*., 1999). Even though there

is no strong evidence, electromagnetic fields have also been hypothesised to affect breast cancer risk by suppressing melatonin production (Kheifets and Matkin, 1999). Other occupational studies among pharmaceutical industry workers, cosmetologists, beauticians, chemists, teachers, social workers, and cashiers have found association with breast cancer risk (Goldberg and Labreche, 1996; Welp *et al*., 1998).

McDowall defined electromagnetic field exposure as living within 30 m of either electrical installation equipment or an overhead power cable (McDowall, 1986). Experimental studies in laboratory animals supports a link between Extremely low-frequency (ELF) magnetic field exposure and decreased melatonin levels but limited data support this link in humans (Reif *et al*., 1995; Davis *et al*., 2006). It has been hypothesized that disruption of the normal nocturnal rise in melatonin resulting from exposure to ELF magnetic fields could increase breast cancer risk (Stevens *et al*., 1992). However, a large number of studies found no evidence that exposure to residential magnetic fields is associated with increased risk (McDowall, 1986; Schreiber *et al*., 1993; Davis *et al*., 2002; London *et al*., 2003). A meta-analysis of 15 case control studies from 2000 to 2009, involving 24,338 cases and 60,628 controls, found no significant association between breast cancer risk in relation to ELF-EMF exposure, even when stratifying by menopausal status or the source of exposure (Chen *et al*., 2010).

2.2.6.3. Trace Elements and Heavy Metals

Trace elements and heavy metals occurring naturally in the environment may influence a woman's risk of developing breast cancer. Exposure to naturally occurring trace elements and heavy metals can be from a variety of sources with geographic variation, including drinking water, air, food, and occupational exposure. Some trace elements such as arsenic (IARC, 1987) and some heavy metals such as cadmium (IARC, 1993) and lead (IARC, 1987) are considered by IARC to be either known or suspected human carcinogens at specified doses of exposure. Others, such as selenium, copper,

iron, and zinc, may plausibly be associated with breast cancer risk given their biological roles. However, evidence associating exposure to these elements with breast cancer risk is limited (Navarro Silvera and Rohan, 2007). Association studies with exposures to heavy metals is also limited and predominantly null, but some studies have found positive association between breast cancer risk and exposure to cadmium (McElroy *et al*., 2006). However, studies on the carcinogenicity of selenium and arsenic does not provide clear evidence to support that they increase breast cancer risk (Garland *et al*., 1996; Navarro Silvera and Rohan, 2007). Copper and iron are biologically important in the production of reactive oxygen species, but excessive exposure to these metals could contribute to oxidative stress and, potentially carcinogenesis. But studies do not give proper association between breast cancer risk and these two metals.

2.3. Mitochondrial DNA

The mitochondrial genome though small is responsible for ensuring that the powerhouses of our cells function properly. As a by-product of their role as powerhouses of our cells, mitochondria generate reactive oxygen species (ROS). This circular genome is more in quantity than its nuclear counterpart and has a higher mutation rate than the nuclear genome and represents less than 1 percent of total cellular DNA. Mammalian mtDNA contains no introns and lacks histones. This, along with its close proximity to ROS produced through oxidative phosphorylation in the mitochondria, make mtDNA vulnerable to oxidative damage and mutations. In fact, the mutation frequency in mtDNA is approximately tenfold greater than that in nuclear DNA (Johns, 1995; Grossman and Shoubridge, 1996). Mutations in the mitochondrial genome have been associated with diverse forms of human disease. Over the years, a probable role for mitochondria in both aging and cancer has emerged. ROS production has been proposed to cause somatic mitochondrial mutations. A large body of evidence suggests that somatic mtDNA mutations play a role in breast carcinogenesis.

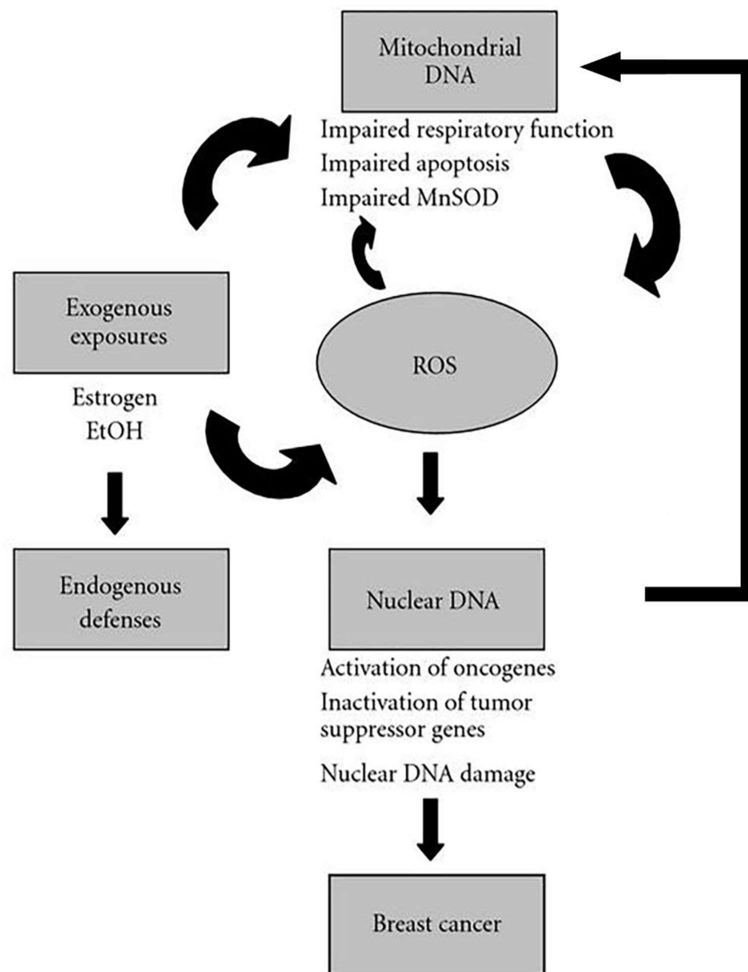


Figure 8. Schema showing how ROS may affect mitochondrial and nuclear DNA leading to breast carcinogenesis (Rohan *et al*., 2010).

2.3.1. Displacement Loop

The d-loop region has been shown to be a mutational 'hot spot' in human cancer. The incidence of somatic mutations in the d-loop region is found in all tumors examined to date and appears to be a universal feature of all cancers (Modica-Napolitano *et al*., 2007). Studies show somatic mutations in majority of breast cancer patients and most of the mutations identified were in the d-loop region (Parrella *et al*., 2001; Tan *et al*., 2002). In a cohort study focusing on the d-loop, 36.36 percent of samples presented somatic mutations while 90.91 percent of samples showed germline mutations (Barekati *et al*.,

2010). Within this region, a poly-C repeat stretch, named D310, contained the majority of mutations (Santos *et al*., 2012; Xu *et al*., 2012). D310 alterations were more frequent in cervical cancer followed by bladder cancer, breast cancer and endometrial cancer (Parrella *et al*., 2003). Among the Chinese, the germline polymorphism of T16189C is suggested to convey increased risk considering the high frequency observed in breast cancer patients (Wang *et al*., 2006). A study among non-Jewish European American found variants 12308G and 10398G to increase breast cancer risk (Covarrubias *et al*., 2008). In another study d-loop mutations were associated with advanced age (>50 years), negative estrogen and progesterone receptor status, as well as poorer disease-free survival (Tseng *et al*., 2006). A study from China of D310 mutations in familial breast cancer recorded extremely high frequencies (Yu *et al*., 2008). The first study on breast cancer and association with mtDNA among the Mizos found mitochondrial gene alterations may attribute for risk (Ghatak *et al*., 2014).

2.3.2. Cytochrome C Oxidase

Cytochrome c oxidase is a large integral membrane protein which is encoded in the mitochondrial genome. It is a terminal oxidase of the mitochondrial electron transport chain, and is expressed in the mitochondrial inner membrane. It is responsible for production of a critical enzyme that controls mitochondrial respiration and is central to apoptosis (Payne *et al*., 2005). This enzyme is the terminal electron acceptor in the electron-transport chain and catalyses the complete reduction of molecular oxygen to water with the supply of four electrons from cytochrome c and four protons taken up from the mitochondrial matrix. Reactive oxygen species (ROS) are known for transient existence, induce oxidative damage leading to both nuclear DNA and mtDNA aberrations, and thus play an important role in carcinogenesis. Increased ROS generation may alter signal transduction pathways, resulting in activation of oncogenes or inactivation of tumor suppressor genes. Defects in the mitochondrial genome and function are suspected to contribute to the development and progression of cancer (Ye

et al., 2008). A depletion of COI subunit has been observed in breast cancer (Putignani *et al*., 2008). Several cancer associated mutations found in mtDNA result in structural modifications of cytochrome c oxidase (Namslauer and Brzezinksi, 2009).

2.4.2. Glutathione S-transferases

In 1961, an enzymatic reaction responsible for the first step in the conjugation of xenobiotics with glutathione was recognized (Combes and Stakelum, 1961). Glutathione s-transferases (GSTs) represent an important group of enzymes encoded by a superfamily of GST genes that detoxify both endogenous compounds and foreign chemicals such as pharmaceuticals and environmental pollutants. Allelic variations are found in genes encoding for these GSTs (Board *et al.*, 1997; Blackburn *et al*., 2000; Strange *et al*., 2000). In estrogen metabolism, GSTs play a role in the catalysis of glutathione (GSH) conjugation of catechol estrogen quinones, the reactive intermediates of estrogen metabolism capable of binding to DNA (Raftogianis *et al*., 2000). GSTs may also be involved in the activation of some carcinogens such as halogenated hydrocarbons (Taningher *et al*., 1999; Strange *et al*., 2000). Human tissues show differential expression of the multiple forms of GSTs (Johansson and Mannervik, 2001). The absence of specific isoenzymes affects the tolerance of organisms to chemical challenges and may result in an increased rate of somatic mutations and higher susceptibility to disease. The ability of many tumours to exhibit increased levels of intracellular GST expression has been linked to mechanisms of chemotherapeutic drug resistance (Black *et al*., 1990). Carriers of homozygous deletions in the GSTM1 and GSTTI genes have an absence of GST mu and GST theta enzyme activity, respectively (Seidegard *et al*., 1988, Hallier *et al*., 1993, Pemble *et al*., 1994). These deletion variants have been useful for molecular epidemiological studies of cancer because they divide study subjects into two well-defined susceptibility classes: those who are and those who are not able to detoxify potential carcinogens by the metabolic pathways regulated by GSTM1 and GSTTI. Studies among the Mizos

indicate GSTM1 and GSTT1 null genotypes were associated with an increased risk of breast cancer (Kimi *et al*., 2016).

2.4.1. Glutathione S-transferase Mu (GSTM1)

GSTM1 is located in the middle of a cluster of five mu class genes on chromosome 1p13.3 (Pearson *et al*., 1993). The homozygous deletion (null genotype) of the GSTM1 gene leads to the total absence of the respective enzyme activity (Seidegard *et al*., 1988). The frequency of the null genotype is around 50 percent in Caucasians and Asians, but only 27 percent in Africans (Garte *et al*., 2001). Allelic variants of GSTM1, A and B have also been found to exist, but based on current knowledge they have no consequences in the catalytic activity of the enzyme (Widersten *et al*., 1991). In addition, gene duplication has been found to exist in Saudi Arabians (McLellan *et al*., 1997). The GSTM1 genotype has been examined in relation to individual breast cancer risk in several studies. Some of these studies pointed to an association between GSTM1 null genotype and breast cancer risk in postmenopausal women (Helzlsouer *et al*., 1998; Charrier *et al*., 1999). But a large number of studies conducted did not find any link between GSTM1 null genotype with breast cancer risk (Zhong *et al*., 1993; Ambrosone *et al*., 1995; Kelsey *et al*., 1997; Bailey *et al*., 1998; Ambrosone *et al*., 1999; Millikan *et al*., 2000). An increased risk for premenopausal women has also been shown, but only in one study (Park *et al*., 2000). Despite these discrepant findings, one meta-analysis suggested that the GSTM1 null genotype poses a moderately increased risk for postmenopausal breast cancer (Dunning *et al*., 1999). Furthermore, the risk has been shown to be modified by BMI (Helzlsouer *et al*., 1998), family history (Millikan *et al*., 2000), use of alcohol (Park *et al*., 2000), and smoking (Millikan *et al*., 2000). The GSTM1 null genotype was not found to be associated with tumour characteristics or survival in one study (Lizard-Nacol *et al*., 1999) but has been suggested to be associated with both longer (Kelsey *et al*., 1997) and shorter (Nedelcheva Kristensen *et al*., 1998) survival. In the study showing a reduced survival

time, the concurrent presence of both the GSTM1 null and GSTT1 null genotypes was associated with positive lymph node status. A strong association has also been shown between GSTM1 deletion and increased PAH-DNA adduct levels in breast tumour tissue (Rundle *et al*., 2000).

2.4.2. Glutathione S-transferase Theta (GSTT1)

The human GSTT1 gene is localised on chromosome 22q with the GSTT2 gene and a pseudogene (Pemble *et al*., 1994; Tan *et al*., 1995). Like GSTM1, homozygous deletion of the GSTT1 gene, leading to total absence of the respective enzyme activity has been observed. Large inter-ethnic differences have been reported in the frequencies of the GSTT1 null genotype being significantly lower among Caucasians (20%) compared to Asians (60%) (Nelson *et al*., 1995). GSTT1 is expressed in human erythrocytes, and various tissues including liver but no expression in breast tissue has been reported (Pemble *et al*., 1994; Landi, 2000). In most of the studies on GSTT1 genotype and breast cancer risk, no significant association has been found (Bailey *et al*., 1998; Helzlsouer *et al*., 1998; Curran *et al*., 2000; Millikan *et al*., 2000; Park *et al*., 2000). However, in one study the risk was found to be modified by the use of alcohol (Helzlsouer *et al*., 1998), and in another study a remarkably lower risk was suggested for premenopausal women lacking the GSTT1 gene (Garcia-Closas *et al*., 1999).

There are only a few published data on breast cancer among the Mizo population. One study reported breast cancer to be cause by betel quid use (Kaushal *et al*., 2010), and another on GSTM1 and GSTT1 null genotypes (Kimi *et al*., 2016). It was also reported that mitochondrial gene alterations may attribute breast cancer risk (Ghatak *et al*., 2014). Some of the indigenous foods of the Mizos contain ash filtrate (*ching-al*), smoked or sun-dried or fermented meats and vegetables. Studies on the association of stomach cancer risk with some of the traditional food reveals elevated risk (Phukan *et al*., 2006, 2014); but there are no data available to associate or even to dissociate. Like a study in Japan showed no association of fermented soyabean (bekang – a delicacies

of the Mizos) with risk (Shirabe *et al*., 2021). With breast cancer incidence on the rise and with only a few data available, it is of great importance to identify the specific risk factors prevalent in this region. The Mizos with their unique lifestyle and dietary habits, the reports from other parts of the world are sometimes not applicable or sufficient, or sometimes not acceptable. Since breast cancer etiology is multifactorial, and differs geographically and ethnically, the question remain what could be the rise in incidence. Is there any change that was not there 50 years or so before. With this in mind, this study was designed to include genes related to catabolism and detoxification of xenobiotics, any mitochondrial gene alterations and any regional characteristics that could influence risk; or whether risk increase when there is association of gene alterations with any risk factors.

CHAPTER 3

OBJECTIVES

3. Objectives

i) To determine the potential demographic risk factors associated with breast cancer in Mizo population.

ii) Study of mitochondrial control region (D-loop), cytochrome c oxidase subunit I (CO1) sequence variations associated with breast cancer risk.

iii) Genes related to catabolism and detoxification of xenobiotics (GSTM1 and GSTT1) to explore their contribution for breast cancer.

CHAPTER 4

MATERIALS AND METHODS

4. Materials and Methods

4.1. Data Source and Sample Collection

A retrospective case control study was conducted with the approval of the Institutional Ethics Committee, Civil Hospital Aizawl [No. B 12018/1/13-CH(A)/IEC/28 of October 15, 2014]. The study was initiated in 2014 with a follow-up after 5 years in 2020. Case data was collected from breast cancer patients registered between 1998-2014, a period spanning for 17 years. Control data was collected from outdoor patient who visited Aizawl Civil Hospital and from personal invitation. Informed consent (Annexure 2) was obtained before each interview as instructed by the Institutional Ethics Committee.

Data was collected from medical records corroborated with a structured questionnaire. Criteria for inclusion in the case study includes confirmed breast cancer registered at Mizoram State Cancer Institute, Zemabawk and Population Based Cancer Registry, Aizawl Civil Hospital; Mizo female with proper contact information for follow-up. All the 758 registered cases were not included in this study because of not meeting the criteria mentioned above, only 363 cases were included. The normal controls were volunteers from outdoor patient who visited Aizawl Civil Hospital and from personal invitation. The criteria for inclusion in the control group was that the volunteer should be a Mizo female at least 18 years of age or above, free of cancer and with no history of cancer in the family, 405 controls matching the criteria mentioned above were included in the study.

4.2. Questionnaire

The questionnaire was structured in such a manner that established as well as suspected risk factors were included (Annexure 3). The questionnaire as mentioned in Table 1 included demography, lifestyle, environmental exposure, medical history, reproductive history, family history and anthropometric factors. The questionnaire

was specifically formed to study education, employment, sleep pattern and duration, physical activity, age at menarche, age at first childbirth, total number of children, duration of breast feeding (in months), height (feet), and weight (kilogram) at the time of breast cancer diagnosis or for controls at the time of interview. Both cases and controls were subjected to the same questionnaire format except for questions about breast cancer. Residence of 5 years in an area before having breast cancer for cases was taken as residence even if they had moved to other areas during interview. The interview was conducted in the local language (Annexure 3) and took about 30-40 minutes for each volunteer. Open-ended question format was used and later categorized. This kind of format does not require a response based on a specific list of choices and allows the participants to answer freely without fear of being wrong. To facilitate comparison, variables were later categorized in two or three or four. For habits, 'quit' is when they have stopped the habit for at least 5 years, for cases before diagnosis and control before interview. Volunteers were recorded as being passive smokers only when there was constant known exposure to secondhand smoke either at work place or at home.

The questions relating to dietary composition and nutrition had a short coming as the subjects were susceptible to both recall bias and selection bias especially against fondness of the food. As such, the questionnaire was structured on frequency of consumption rather than amount which could still be influenced by selection bias especially against fondness of the food. But they do provide a rough idea of their daily dietary intake, but not a prolonged accurate dietary habit. Less is when consumption frequency is trice or less in a month; moderate when they consume once to four times in a week; high is when they consume it at least five times or more in a week.

Total time of breastfeeding was calculated as mean time of breastfeeding multiplied with parity. To calculate mean duration total time of breastfeeding was divided by number of children. Job and age gap between children was also considered but 15 or 18 months given if response is more than a year, unless specific mention of

Table 1. Interview questionnaires.

Category	**Variables**
Personal information	Name, age, residence, occupation, education, handedness, weight, height
Lifestyle	Behavioral habits (including secondhand smoke), diet, sleep, physical activity
Environmental exposure	Electromagnetic and pesticide exposures
Medical history	All major illness including diabetes mellitus, hypertension, any cancer
Menstrual history	Menarche and menopause ages, dysmenorrhea, menopausal problems
Reproductive history	Marriage, parity, breastfeeding history, use of contraceptives
Family history	Breast cancer history (1st degree, 2nd degree and 3rd degree), maternal age
Knowledge	Self-examination, clinical examination, breast cancer-related knowledge

time (Butt *et al*., 2014). All live births were included in the study as death of a child after birth had to be considered in some of the volunteers.

4.3. Anthropometric and Physical Activity Measurement

Body Surface Area was calculated using Mosteller's formula which is the square root of weight (kilogram) times the height (centimeter) divided by 3600. Body Mass Index was measured as weight in kilograms (kg) divided by the square of height in meters (m^2). Although an imperfect measure, BMI is highly correlated with percentage of body fat (Deurenberg *et al*., 1991). The World Health Organization has defined the following cut points for BMI: BMI less than 18.50 is considered underweight; BMI between 18.50 and 24.99 is described as normal or healthy; BMI between 25.00 and 29.99 is grade 1 overweight or overweight; BMI between 30.00 and 39.99 is grade 2 overweight or obese; BMI greater than or equal to 40.00 is grade 3 overweight or morbidly obese. For our study, quartile distributions were used to categorize BMI to facilitate comparison; underweight = BMI less than 18.50; healthy = 18.50-24.99; overweight = 25.00-29.99; obese = 30.00-39.99 (WHO, 1995).

To quantify physical activity, metabolic equivalent of energy expenditure (MET) value was assigned to each reported activity according to the Framingham Physical Activity Index score (Dorgan *et al*., 1994). MET was calculated by summing the number of hours spent in each activity intensity level and multiplying by a respective weight factor derived from the estimated oxygen consumption requirement for each intensity level. One MET being equivalent to the amount of energy a person expends at rest. In this study, MET value was measured for a typical 24-hour duration, categorizing by intensity with slight modification. Since sleep duration was recorded during interview, score of 1 MET was multiplied with the hours spent sleeping or at rest as shown in Table 2. Depending on their main occupation, age and health condition, 5 or 7 hours was multiplied for 'sedentary' for all subjects, the rest of the 24 hours was divided between their main occupation, recreational and household activities. Even though farming is categorized under 'heavy', the number of hours spend for the category differs depending on their age and health condition. If the subject was healthy and under 50 years of age, 5 hours was multiplied. Even if subject was under 50, 2 hours multiplied under 'heavy' if they were occasional farmers. If aged between 51 and 65 years, 2 hours was multiplied, and 1 hour if between age 66 and 69 years, and none if 70 years or more. Physical activity was dichotomized into normal and heavy, where light and moderate MET was grouped as normal; and heavy as heavy to enable further analysis.

Table 2. Physical activity of 24-hour duration.

Sleep (1 x hours)	Sedentary (1.1 x hours)	Slight (1.5 x hours)	Moderate (2.4 x hours)	Heavy (5 x hours)
	Personal time	Knitting	Household work	Farming
	Meal time	Desk job/Teaching	Carpentry	Sports
	Watching television	Shopkeeping	Nursing/beautician	
	Riding in a car	Tailoring	Walking/ Gardening	
	Quite time/meditation	Handloom weavers	Pig/poultry/cattle farm	

4.4. Molecular Study

For molecular study, 49 cases and 41 controls were selected. Healthy controls were individually matched with cases by age (cases = 48.96±11.33, control = 46.73±10.21) except for a 78-year-old female case. Matching area status of incidence, (cases urban=26, rural=23; control urban=25, rural=16) could not be done in spite of inviting residents of Ramlaitui, a rural area of Lunglei District. But to at least match standard of living like farming, residents of Tlangnuam, Aizawl District were personally invited to volunteer for the study (Physical activity score of cases = 36.22±3.12, control = 36.17±2.41; BMI of cases = 23.61±3.65, control = 23.12±3.16).

4.4.1. DNA Isolation

Signed informed consent was obtained from all subjects prior to collection of blood samples. Peripheral blood sample was collected in a 2 ml EDTA vacuum tube and stored at -20ºC. DNA isolation from blood was done using a commercially available kit from Qiagen (Qiagen DNeasy Blood & Tissue Kit, Cat. No.69504). Isolation was done as per instructions given along with the kit. Into a 1.75 ml microcentrifuge tube, 20 µl of proteinase K and 150 µl of anticoagulated blood was added. The volume was adjusted to 200 µl with PBS (pH 7.4). The sample vial was inverted before pipetting out blood. Buffer AL (without added ethanol) of 200 µl was added and mixed thoroughly by vortexing for 10-20 seconds, centrifuged at 8000 rpm for 1 minute so that no solution sticks on the cap. After the sample vial was incubated at 56°C for 10 minutes, 200 µl ethanol (96–100%) was added, and mixed thoroughly by vortex and centrifuge. Lysate was transferred in DNeasy mini spin column placed in a 2 ml collection tube without wetting the rim. Centrifuged at 8000 rpm for 1 min. Flow-through and collection tube was discarded. The DNeasy mini spin column was placed in a new 2 ml collection tube and 500 µl Buffer AW1 was added. Centrifuged for 1 min at 8000 rpm, flow-through and collection tube discarded. The DNeasy mini spin column was placed in a new 2 ml collection tube and 500 µl Buffer AW2 was added.

Centrifuged for 3 min at 14,000 rpm to dry the DNeasy membrane. Flow-through and collection tube discarded. The DNeasy mini spin column was placed in a clean 1.5 ml microcentrifuge tube. Buffer AE of 200 μl was added directly onto the DNeasy membrane. Incubated at room temperature for 1 min, and then centrifuged for 1 min at 8000 rpm to elute. The extracted DNA was stored in at -20ºC until further use.

4.4.2. Amplification of Gene

Table 3. List of primers and sequences.

Gene	**Primer Sequence 5′ – 3′**
D-loop	FW: AAGACTCGGCAGCATCTCCACACCATTAGCACCCAAAGCT RW: GCGATCGTCACTGTTCTCCACTGTTAAAAGTGCATACCGCCA
CO1	FW: AAGACTCGGCAGCATCTCCATCAACAAATCATAAAGATATTGG RW: GCGATCGTCACTGTTCTCCAGGGTGACCAAAAAATC
GSTT1	FW: TTCCTTACTGGTCCTCACATCTC RW: TCACCGGATCATGGCCAGCA
GSTM1	FW: GAACTCCCTGAAAAGCTAAAGC RW: GTTGGGCTCAAATATACGGTGG
Alu	FW: GCCCTCTGCTAACAAGTCCTAC RW: GCCCTAAAAAGAAAATCGCCAAT

D-loop

The mtDNA d-loop region was amplified by PCR using forward 5′-AAGACTCGGCAGCATCTCCACACCATTAGCACCCAAAGCT-3′ and reverse 5′-GCGATCGTCACTGTTCTCCACTGTTAAAAGTGCATACCGCCA-3′ as shown in Table 3. Amplification was performed on VeritiDx, a thermal cycler from Thermo Fisher Scientific using. Primers were used at a final concentration of 200 nM and dNTPs at 0.2 mM; MgCl2 was used at a final concentration of 1.5 mM. An enzyme blend of Platinum taq (Invitrogen, Cat. No. 10966034) and Hotstar taq (Qiagen, Cat. No. 1007837) was used. The reaction mixture was heated to 95 °C for 15 min, followed by 35 cycles each consisting of 1 min denaturation at 95 °C, 1 min annealing at 60 °C, 90 sec of extension at 72 °C and a final 10 min extension at 72 °C as shown in Table

4. The amplified product was purified and sequenced by next generation sequencing of targeted amplicon synthesis chemistry (Illumina MiSeq). Bioinformatic analysis was performed using Illumina's cloud computing platform, Basespace.

Table 4. Thermal profile for D-Loop.

Stage	Steps	Temperature (°C)	Time	No. of Cycles
Stage 1	Initial denaturation	95°C	15 minutes	Hold
Stage 2	Denaturation	95°C	1 minutes	35 cycles
	Annealing	60°C	1 minutes	
	Extension	72°C	90 seconds	
Stage 3	Final Extension	72°C	10 minutes	Hold
Stage 4	Hold	4°C	infinite	Hold

CO1

The mtDNA CO1 region was amplified by PCR using forward 5′-AAGACTCGGCAGCATCTCCATCAACAAATCATAAAGATATTGG-3′ and reverse 5′-GCGATCGTCACTGTTCTCCAGGGTGACCAAAAAATC-3′ as shown in Table 3. Amplification was performed on VeritiDx, a thermal cycler from Thermo

Table 5. Thermal profile for CO1.

Stage	Steps	Temperature (°C)	Time	No. of Cycles
Stage 1	Initial denaturation	95°C	15 minutes	Hold
Stage 2	Denaturation	95°C	30 seconds	35 cycles
	Annealing	52°C	30 seconds	
	Extension	72°C	90 seconds	
Stage 3	Final Extension	72°C	10 minutes	Hold
Stage 4	Hold	4°C	infinite	Hold

Fisher Scientific. Primers were used at a final concentration of 200 nM and dNTPs at 0.2 mM; MgCl2 was used at a final concentration of 2.5 mM. An enzyme blend of Platinum taq (Invitrogen, Cat. No. 10966034) and Hotstar taq (Qiagen, Cat. No. 1007837) was used. The reaction mixture was heated to 95 °C for 15 min, followed by 35 cycles each consisting of 30 sec denaturation at 95 °C, 30 sec annealing at 52 °C, 90 sec of extension at 72 °C and a final 10 min extension at 72 °C as shown in Table 5. The amplified product was purified and sequenced by next generation sequencing of targeted amplicon synthesis chemistry (Illumina MiSeq). Bioinformatic analysis was performed using Illumina's cloud computing platform, Basespace.

Glutathione s-transferase

GSTT1 was amplified by PCR using primers as forward 5′-TTCCTTACTGGTCCTCACATCTC-3′ and reverse 5′-TCACCGGATCATGGCC AGCA -3′; GSTM1 primers as forward 5′-GAACTCCCTGAAAAGCTAAAGC-3′ and reverse 5′-GTTGGGCTCAAATATACGGTGG-3′; and a positive control Alu as forward 5′-GCCCTCTGCTAACAAGTCCTAC-3′ and reverse 5′-GCCCTAA AAAGAAAATCGCCAAT-3′ as shown in Table 3. The PCR conditions for the GSTM, GSTT, and the control gene (*Alu*) were standardized and analysis for these genes was performed in a multiplex PCR according to previously described methods (Arand *et al.,* 1996). The PCR was performed using the QuantiTect 2x mastermix (Qiagen Cat. No. 204543), 200 nM each primer, and approximately 50 ng of DNA extracted from the blood samples provided. The reaction mixture was heated to 95 °C for 15 min, followed by 35 cycles each consisting of 20 sec denaturation at 95 °C, 25 sec annealing at 63 °C, 35 sec of extension at 72 °C and a final 10 min extension at 72 °C as shown in Table 6. The total reaction volume was 20µl; 2 µl of the amplified product was analyzed on a 2% agarose gel which resulted in a 215 bp fragment for GSTM1, 480 bp fragment for GSTT1 and 350 bp fragment for albumin gene as an internal control (Egan *et al*., 2004). The absence of the specific GSTM1 and/or GSTT1

Table 6. Thermal profile for multiplex PCR.

Stage	Steps	Temperature (°C)	Time	No. of Cycles
Stage 1	Initial denaturation	95°C	15 minutes	Hold
Stage 2	Denaturation	95°C	20 seconds	35 cycles
	Annealing	63°C	25seconds	
	Extension	72°C	35 seconds	
Stage 3	Final Extension	72°C	10 minutes	Hold
Stage 4	Hold	4°C	infinite	Hold

fragments specify the parallel null genotype, whereas the presence of the albumin gene fragment confirms that the accepted null genotype was not due to PCR failure. A No Template Control (NTC), Positive Control (PC) that was positive for all three genes, and FlashGel marker (Lonza) was run with each batch of samples (Figure 9).

4.5. DNA Analysis

D-loop and CO1

Sequencing of D-loop and CO1 was obtained using NGS (Illumina MiSeq). Genomic visualization was done using Homo sapiens (Human) GRCh37 hg19 (Feb 2009) as reference on GenomeBrowse 3.0.0 and variant analysis was done using VarSeq 1.5.0 (Golden Helix 1.5.0). Variant pathogenicity was done using HmtVar. HmtVar is a manually curated database offering variability and pathogenicity information about mtDNA variants. Data were gathered from HmtVar's twin database HmtDB, and further integrated with pathogenicity predictions as well as additional information from several online resources focused on mtDNA, such as MITOMAP, 1000 Genomes Project, MutPred, SNPs and GO and many others. Sequences were also evaluated against the Mitomaster (https://www.mitomap.org/) for variants and haplogroup. Mitomaster uses HaploGrep2 with Phylotree 17 for haplogroup determination NC_012920.1.

Phylogenetic analysis was performed using the sequences of d-loop and CO1 of all the samples, an outgroup sequence (accession no NC_012920.1) was selected and retrieved from NCBI. All the sequences were aligned using clustalW implemented in Mega 11 (Tamura *et al.*, 2021) and phylogenetic tree was build using maximum likelihood (ML) and neighbourhood joining (NJ), default parameter was used. Phylip file was generated using ALTER Alignment Transformation Environment (Glez-Peña *et al*., 2010) for RaxML analysis. Maximum likelihood of the sequences was then generated using the program raxmlGUI 2.0 (Edler *et al*., 2021), bootstrap value was set at 500, substitution matrix and substitution rates was set at GTR and GAMMA.

Glutathione s-transferase

The presence of GSTT1 and GSTM1 fragments confirmed that the sample was of normal genotype. The presence or absence of fragments can be visualized as shown in Figure 9 in the results chapter. Whereas the absence of the GSTT1 and GSTM1 fragments confirmed that the sample was of null genotype. GSTT1 and GSTM1 were combined as having no deletions, wild type GSTT1 with null GSTM1, null GSTT1 with wild type GSTM1 or having homozygous deletions.

4.6. Statistical Methods

Risk factors were compared between cases and controls by constructing frequency distribution using IBM SPSS Statistics 20.0. To test the significance of differences between two groups, Pearson's chi-square test was adopted. Variables were further categorized into 8 groups - demography, dietary habits, behavioral habits, environmental exposure, medical history, menstrual history, reproductive history and family history. Variables with a P value less than 0.05 were observed as relevant to show significant differences between case and control groups.

$$x^2 = \sum \frac{(O - E)^2}{E}$$

O = Observed frequency; E = Expected frequency

To examine the causal effects of different risk factors, univariate and multivariate logistic regression were used. Odd ratio (ORs) with 95% confidence intervals (CIs) of case and control groups were calculated from these regression estimates to interpret the severity of each factor. In addition to this, survival analysis of the cases was also conducted using Cox-regression model. Subjects with missing values in any of the variables in a regression model were excluded from the analysis. Variables in which when both case and control have the same value were also excluded (Yu *et al*., 2012).

Statistical test was also performed on a few factors based on menopausal status. A woman was considered postmenopausal: if she had undergone bilateral oophorectomy; if she affirmed that her menstruations had ceased for at least 6 consecutive months before diagnosis of breast cancer or prior to interview for control; if the above information were unavailable or inconclusive, 55 years and above was considered as postmenopausal for the study. To facilitate comparison, variables were categorized in two, never and ever or low and high. To examine the causal effects of the selected factor, univariate and multivariate logistic regression were used. BMI (kg/m^2) were dichotomized based on the median values (22.49) of controls.

Association of risk factors with the selected genes was also performed. To calculate the population characteristics and test differences between molecular study group chi-square was used where appropriate. Variables were also categorized into

two where possible. To examine the causal effects of different factors, univariate and multivariate logistic regression were used. Odd ratio (ORs) with 95% confidence intervals (CIs) of case and control groups were calculated from these regression estimates to interpret the severity of each factor. Normal or wild type GSTT1 and GSTM1 genotypes and having no deletion in GSTT1 and GSTM1 were used as referent. The adjusted ORs were evaluated by including the potential confounders such as age, residence, educational qualification and parity.

CHAPTER 5

RESULTS

5. Results

5.1. Demographic Characteristics

Among confirmed breast cancer cases, 749 were females with a mean age of 49.71 (SD = 12.12) and 9 males with a mean age of 59.89 (SD = 17.92). Of the 758 cases, with an age range of 21-91 years (Males = 9; non-Mizo = 18 including 2 non-Mizo males) only 376 cases were interviewed because contact information was either lost (378 cases) and 4 declined an interview. Healthy controls of 405 volunteers were with an age range of 18-71. For epidemiological analysis, 363 cases and 405 controls were compared. For molecular analysis, 49 cases and 41 controls were studied.

Table 7. Age distribution of cases (1998-2014).

Age Distribution	Female	Male	No. of cases
≤29	22	0	22
30-34	36	0	36
35-39	91	2	93
40-44	116	1	117
45-49	129	0	129
50-54	115	0	115
55-59	75	1	76
60-64	63	1	64
65-69	46	1	47
70-74	33	1	34
≥75	23	2	25
Total	**749**	**9**	**758**

Table 8. Breast cancer incidence (1998-2014) districtwise.
(Aizawl Municipal Corporation area as urban)

District	**No. of cases**
Aizawl	438 (Urban=351|Rural=87)
Champhai	80
Kolasib	45
Lawngtlai	22
Lunglei	72
Mamit	31
Saiha	18
Serchhip	37
Others	15 (Myanmar=12|Tripura=3)
Total	**758**

Table 9. Number of cases registered and interviewed.

Year	Number of cases registered			No. of cases interviewed
	MSCI	PBCR	Total	
1998	1	-	1	1
1999	7	-	7	2
2000	14	-	14	2
2001	6	-	6	-
2002	17	1	18	2
2003	26	24	50	7
2004	36	21	57	12
2005	31	11	42	11
2006	37	16	53	13
2007	35	23	58	16
2008	44	5	49	23
2009	45	14	59	28
2010	41	10	51	28
2011	54	32	86	56
2012	52	13	65	51
2013	68	12	80	69
2014	62	-	62	55
Total	**576**	**182**	**758**	**376**

Table 10. Comparison of mean and standard deviation between groups.
(case = 363; control = 405)

Variables	Cases ± SD	Control ± SD	Total ± SD
Age	49.64 ± 11.87	38.82 ± 12.39	43.93± 13.29
Height (cm)	153.72 ± 5.57	154.90 ± 4.34	154.34 ± 4.99
Weight (kg)	54.75 ± 9.64	53.76 ± 8.41	54.23 ± 9.02
BSA	1.52 ± 0.15	1.52 ± 0.13	1.52 ± 0.14
BMI	23.15 ± 3.70	22.49 ± 3.35	22.80 ± 3.53
Physical activity (MET)	36.72 ± 3.08	35.83 ± 2.31	36.25 ± 2.74
Maternal age	28.34 ± 6.17	28.47 ± 7.15	28.41 ± 6.70
First child birth[*1]	23.89 ± 5.04	22.23 ± 4.96	23.06 ± 5.06
Last child birth[*1]	31.32 ± 5.88	29.25 ± 5.62	30.29 ± 5.84
Parity[*1]	3.63 ± 1.73	3.34 ± 1.84	3.49 ± 1.79
Age at menarche	15.11 ± 1.43	14.81 ± 1.49	14.95 ± 1.47
Age at menopause[*2]	48.44 ± 4.66	47.21 ± 4.98	48.00 ± 4.80
Menstrual duration[*2]	33.20 ± 4.92	32.15 ± 4.99	32.82 ± 4.94

([*1] Ca = 298, Co =297; [*2] Ca = 177, Co =99)

Most respondents were subjects themselves (cases = 286, controls = 405). But, for 77 cases, respondents were their primary caregivers who had reliable information like either their sibling (n=22), husband (n=21), children (n=20), mother (n=4), close relative (n=7) and volunteer caregiver (n=3) as they were too ill and some had passed away. In spite of this, all respondents could provide reliable information. Initially, medications like pain-relief drugs, hormone replacements, hypertensive drugs, anti-diabetic drugs were included in the questionnaire, but most of the respondents could not recall type of medicine used especially for menopausal problems, but for contraceptive pills, most of them could provide at least the brand name.

The tumor histology of the breast cancer group was 95.4% Invasive Ductal Carcinoma (IDC), 3.2% Ductal Carcinomain Situ (DCIS), 1.1% Invasive Lobular Carcinoma (ILC) and 0.3% Lobular Carcinoma in Situ (LCIS). The molecular sub-

type was 60.6% Luminal A, 12.4% Luminal B, 11.5% with Her2 Overexpression and 15.5% Triple Negative (Table 11). Data of tumor histology and molecular sub-type was missing for 15 cases.

Table 11. Tumor histology and molecular sub-types of breast cancer cases.

Tm Histology	**n** (%)		**Molecular Sub-type**	**n** (%)
Invasive Ductal Carcinoma	332 (95.4)		Luminal A	211 (60.6)
Invasive Lobular Carcinoma	4 (1.1)		Luminal B	43 (12.4)
Ductal Carcinomain Situ	11 (3.2)		Her2 Overexpression	40 (11.5)
Lobular Carcinoma in Situ	1 (0.3)		Triple Negative	54 (15.5)

Molecular Sub-type	IDC	ILC	DCIS	LCIS
Luminal A	200	3	7	1
Luminal B	42	1	-	-
Her2 Overexpression	38	-	2	-
Triple Negative	52	-	2	-

5.2. Comparison of Variables

5.2.1. Demography

Overall area of incidence for urban and rural is 351 and 407 respectively. But 185 cases from urban areas and 178 cases from rural areas were included in this study. Contact information was lost for 229 cases from rural area, and 162 cases from urban areas with 4 declining to participate. No differences were found between cases and controls with regards to education and handedness. There was a significant difference between cases and controls for residence, employment status, BMI, marital status, age at marriage, number of siblings and their maternal age at their birth (Table 12).

5.2.2. Diet and Nutrition

There was no significant difference in salt intake and consumption of pork. Significant differences were observed between case and control groups in their water intake, consumption of sodium bicarbonate, fruits, vegetables, red meat, poultry, sea food and fermented pork (Table 13).

Table 12. Comparison of demographic characteristics between case and control.

Variables	Cases **n** (%)	Controls **n** (%)	χ^2 (df)	*P* values
Residence			47.005 (1)	0.000
Urban	185 (51.0)	303 (74.8)		
Rural	178 (49.0)	102 (25.2)		
Education			5.704 (3)	0.127
Illiterate	24 (6.6)	12 (3.0)		
Primary-Middle	181 (49.9)	210 (51.9)		
High-12	127 (35.0)	147 (36.30)		
≥Graduate	31 (8.5)	36 (8.8)		
Employment status			56.337 (3)	0.000
Housewife	59 (16.3)	162 (40.0)		
Manual Worker	140 (38.6)	132 (32.6)		
Non-Manual	92 (25.3)	63 (15.6)		
Employed/r	72 (19.8)	48 (11.8)		
BMI			96.758 (3)	0.000
Underweight	24 (6.6)	18 (4.4)		
Normal	195 (53.7)	96 (23.7)		
Overweight	128 (35.3)	285 (70.4)		
Obese	16 (4.4)	6 (1.5)		
Handedness			0.432 (2)	0.806
Left	43 (11.9)	42 (10.4)		
Right	309 (85.1)	351 (86.7)		
Both	11 (3.0)	12 (2.9)		
Marital status			13.981 (3)	0.003
Single	54 (14.9)	75 (18.5)		
Married	261 (71.9)	288 (71.1)		
Divorced	14 (3.9)	27 (6.7)		
Widow	34 (9.4)	15 (3.7)		
Age at marriage			19.413 (3)	0.000
<19	77 (21.2)	132 (32.6)		
20-34	216 (59.5)	189 (46.7)		
35>	16 (4.4)	9 (2.2)		
Single	54 (14.9)	75 (18.5)		
Siblings			30.591 (3)	0.000
Single Child	7 (1.9)	15 (3.7)		
1-2 sibs	30 (8.3)	69 (17.0)		
3-6 sibs	205 (56.5)	246 (60.8)		
>7 sibs	121 (33.3)	75 (18.5)		
Maternal age			10.536 (2)	0.005
<19	10 (2.8)	33 (8.2)		
20-34	285 (78.5)	300 (74.1)		
>35	68 (18.7)	72 (17.7)		

Table 13. Comparison of dietary habits between case and control groups.

Variables	Cases **n** (%)	Controls **n** (%)	χ^2 (df)	*P* values
Water intake			14.515 (3)	0.002
<1L	105 (28.9)	75 (18.5)		
1L/day	146 (40.2)	174 (43.0)		
2L/day	98 (27.0)	126 (31.1)		
3L or more	14 (3.9)	30 (7.4)		
Salt intake			3.122 (2)	0.210
Less	182 (50.2)	201 (49.6)		
Moderate	60 (16.5)	51 (12.6)		
High	121(33.3)	153 (37.8)		
Sodium bicarbonate			42.287 (3)	0.000
None	5 (1.4)	0 (0.0)		
Less	71 (19.6)	156 (38.5)		
Moderate	67 (18.5)	39 (9.6)		
High	220 (60.5)	210 (51.9)		
Fruits			51.381 (3)	0.000
None	27 (7.4)	18 (4.4)		
Less	161 (44.4)	204 (50.4)		
Moderate	88 (24.2)	30 (7.4)		
High	87 (24.0)	153 (37.8)		
Vegetables			12.529 (2)	0.002
Less	2 (0.6)	15 (3.7)		
Moderate	29 (8.0)	18 (4.4)		
High	332 (91.4)	372 (91.9)		
Red meat			7.803 (3)	0.050
None	86 (23.7)	102 (25.2)		
Less	211 (58.1)	258 (63.7)		
Moderate	61 (16.8)	42 (10.4)		
High	5 (1.4)	3 (0.7)		
Poultry			13.785 (3)	0.003
None	24 (6.6)	54 (13.3)		
Less	229 (63.1)	225 (55.6)		
Moderate	94 (25.9)	117 (28.9)		
High	16 (4.4)	9 (2.2)		
Sea food			8.960 (3)	0.030
None	42 (11.6)	78 (19.3)		
Less	239 (65.8)	237 (58.5)		
Moderate	72 (19.8)	78 (19.3)		
High	10 (2.8)	12 (2.9)		
Pork			1.299 (3)	0.729
None	50 (13.8)	60 (14.8)		
Less	121 (33.3)	138 (34.2)		
Moderate	151 (41.6)	171 (42.2)		
High	41 (11.3)	36 (8.8)		
Fermented pork			38.587 (3)	0.000
None	49 (13.5)	57 (14.1)		
Less	247 (68.0)	327 (80.7)		
Moderate	41 (11.3)	6 (1.5)		
High	26 (7.2)	15 (3.7)		

Less = <3/month; **Mod** = 1-4/week; **High** = >5/week-daily (frequency of consumption rather than amount)

Table 14. Comparison of behavioral habits between case and control groups.

Variables	Cases **n** (%)	Controls **n** (%)	χ^2 (df)	*P* values
Betel quid			50.603 (3)	0.000
No	113 (28.9)	60 (14.8)		
Occasional	124 (34.2)	138 (34.1)		
<10/day	81 (22.3)	174 (43.0)		
>10/day	45 (12.4)	33 (8.1)		
Tobacco			6.127 (3)	0.106
No	141 (38.8)	141 (34.8)		
Quit	9 (2.5)	3 (0.7)		
Occasional	20 (5.5)	30 (7.4)		
Regularly	193 (53.2)	231 (57.1)		
Cigarette smoking			38.394 (3)	0.000
No	229 (59.8)	315 (76.3)		
Occasional	37 (10.2)	45 (11.1)		
<10/day	70 (19.3)	42 (10.4)		
>10/day	27 (7.4)	3 (0.7)		
Tuibur			15.803 (3)	0.001
No	250 (68.9)	264 (65.2)		
Quit	7 (1.9)	3 (0.7)		
Occasional	61 (16.8)	108 (26.7)		
Regularly	45 (12.4)	30 (7.4)		
Alcohol			0.265 (2)	0.876
No	344 (94.8)	381 (94.1)		
Occasional	15 (4.1)	18 (4.4)		
Regularly	4 (1.1)	6 (1.5)		
Passive smoking			28.756 (1)	0.000
No	158 (43.5)	102 (25.2)		
Yes	205 (56.5)	303 (74.8)		
Physical activity			13.247 (2)	0.001
Light	75 (20.7)	93 (22.9)		
Moderate	193 (53.2)	249 (61.5)		
Heavy	95 (26.2)	63 (15.6)		
Sleep duration			47.601 (2)	0.000
<5 hrs	35 (9.7)	24 (5.9)		
5-7 hrs	113 (31.1)	51 (12.6)		
>8 hrs	215 (59.2)	330 (81.5)		
Sleep pattern			19.792 (1)	0.000
Undisturbed	217 (59.78)	303 (74.8)		
Disturbed	146 (40.22)	102 (25.2)		

5.2.3. Behavioral Habits

There was no difference between the groups in chewing tobacco and consumption of alcohol. But significant differences were observed in their chewing

of betel quid and *tuibur*, cigarette smoking, passive smoking, physical activity, sleep duration and sleep pattern (Table 14).

5.2.4. Environmental Exposure and Medical History

Residence within 15 feet from electrical and mobile tower, transformer was regarded as proximal and exposed to electromagnetic waves. There was no significant difference between case and control groups with regards to environmental exposure. There was no significant difference between case and control groups with regards to their history of having asthma and allergy and mental disorder. But significant differences were seen in their history of having cancer, diabetes mellitus and hypertension (Table 15).

Table 15. Comparison of environmental exposure and chronic diseases between groups.

Variables	Cases **n** (%)	Controls **n** (%)	χ^2 (df)	*P* values
Pesticide exposure			5.192 (2)	0.075
Unexposed	338 (93.1)	363 (90.0)		
Occasional	6 (1.7)	18 (4.4)		
Exposed	19 (5.2)	24 (5.9)		
EM exposure			0.125 (1)	0.724
Unexposed	257 (70.8)	282 (69.6)		
Exposed	106 (29.2)	123 (30.4)		
History of illness			49.850 (1)	0.000
No	272 (74.9)	378 (93.3)		
Yes	91 (25.1)	27 (6.7)		
Diabetes mellitus			20.683 (1)	0.000
No	322 (88.7)	393 (97.0)		
Yes	41 (11.3)	12 (3.0)		
Hypertension			59.498 (1)	0.000
No	306 (84.3)	402 (99.3)		
Yes	57 (15.7)	3 (0.7)		
Asthma & Allergy			0.232 (1)	0.630
No	350 (96.4)	393 (97.0)		
Yes	13 (3.6)	12 (3.0)		
Cancer			4.486 (1)	0.034
No	359 (98.9)	405 (100)		
Yes	4 (1.1)	0 (0.0)		
Mental disorder			1.117 (1)	0.291
No	362 (99.7)	405 (100)		
Yes	1 (0.3)	0 (0.0)		

Table 16. Comparison of menstrual characteristics between case and control groups.

Variables	Cases **n** (%)	Controls **n** (%)	χ^2 (df)	*P* values
Age at menarche			7.450 (2)	0.024
<12	15 (4.1)	18 (4.4)		
13-15 years	224 (61.7)	285 (70.4)		
>16	124 (34.2)	102 (25.2)		
Menstrual status			49.162 (1)	0.000
Premenopause	186 (51.2)	306 (75.6)		
Postmenopausal	177 (48.8)	99 (24.4)		
Age at menopause			141.295 (3)	0.000
Premenopause	120 (33.1)	306 (75.6)		
<49	129 (35.5)	57 (14.1)		
50-54 years	99 (27.3)	39 (9.6)		
>55	15 (4.1)	3 (0.7)		
Menstrual pattern			0.342 (1)	0.559
Regular	340 (93.7)	375 (92.6)		
Irregular	23 (6.3)	30 (7.4)		
Lifetime menstrual duration			147.024 (2)	0.000
Premenopause	116 (31.9)	306 (75.6)		
<35 years	184 (50.7)	57 (14.1)		
>36 years	63 (17.4)	42 (10.3)		
Dysmenorrhea			122.801 (2)	0.000
No	290 (79.9)	168 (41.5)		
Slight	42 (11.6)	93 (22.9)		
Severe	31 (8.5)	144 (35.6)		
Pain relief meds			35.871 (2)	0.000
No	340 (93.7)	318 (78.5)		
Sometimes	16 (4.4)	57 (14.1)		
Regular	7 (1.9)	30 (7.4)		
Menopausal problem			0.194 (1)	0.660
No	133 (75.1)	72 (72.7)		
Yes	44 (24.9)	27 (27.3)		
Consultation			0.122 (1)	0.726
No	158 (89.3)	87 (87.9)		
Yes	19 (10.7)	12 (12.1)		
Medication			0.661 (1)	0.416
No	161 (91.0)	87 (87.9)		
Yes	16 (9.0)	12 (12.1)		

5.2.5. Menstrual History

No significant differences between case and control groups with regards to menstrual pattern or with having menopausal problems. But significant differences were observed with regards to their age at menarche, menopausal status, age at menopause, lifetime cumulative number of menstrual cycles, dysmenorrhea and pain relief medication for such (Table 16).

5.2.6. Reproductive History

There was no significant difference between case and control groups in the number of miscarriages. But significant differences were observed with parity, breastfeeding and duration of breastfeeding, age at first child birth, age at last child birth, oral contraceptive use and duration of use (Table 17).

5.3. Factors Related to Breast Cancer on Univariate Analysis

Univariate Logistic regression analysis was done on all variables between case and control groups. There were significant association between the groups in the following factors – consumption of sodium bicarbonate, fermented pork, cigarette smoking, physical activity, sleep pattern, diabetes mellitus, hypertension, ages at menarche and menopause, menstrual status, parity, and breastfeeding, age at first child birth and age at last child birth. In this study group, some risk factors seem to have inverse association as in BMI, water intake, betel quid, secondhand smoke, sleep duration, lifetime menstrual cycle, dysmenorrhea and use of medicine to relief the discomfort, oral contraceptive use and duration of use (Table 18).

Table 17. Comparison of reproductive characteristics between case and control.

Variables	Cases **n** (%)	Controls **n** (%)	χ^2 (df)	*P* values
Miscarriage			3.947 (3)	0.267
0	273 (75.2)	297 (73.3)		
1	65 (17.9)	81 (20.0)		
2	17 (4.7)	24 (5.9)		
≥3	8 (2.2)	3 (0.8)		
Parity			12.042 (3)	0.007
0	65 (17.9)	108 (26.7)		
1-2	77 (21.2)	93 (23.0)		
3-4	144 (39.7)	144 (35.5)		
≥5	77 (21.2)	60 (14.8)		
Breastfeeding			6.821 (1)	0.009
Never	68 (18.7)	108 (26.7)		
Ever	295 (81.3)	297 (73.3)		
Breastfeed duration			43.691 (3)	0.000
Never	68 (18.7)	108 (26.7)		
< 6 months	2 (0.6)	6 (1.5)		
7-12 months	64 (17.6)	15 (3.7)		
≥13 months	229 (63.1)	276 (68.1)		
Age at FCB			27.070 (3)	0.000
0	65 (17.9)	108 (26.7)		
< 19	49 (13.5)	93 (23.0)		
20-34	238 (65.6)	192 (47.3)		
> 35	11 (3.0)	12 (3.0)		
Age at LCB			12.926 (3)	0.005
0	65 (17.9)	108 (26.7)		
< 19	4 (1.1)	6 (1.5)		
20-34	209 (57.6)	228 (56.3)		
> 35	85 (23.4)	63 (15.5)		
Oral contraceptive			40.198 (1)	0.000
Never	323 (89.0)	285 (70.4)		
Ever	40 (11.0)	120 (29.6)		
OC duration			40.789 (3)	0.000
Never	323 (89.0)	285 (70.4)		
<1 year	13 (3.6)	42 (10.4)		
1-4 years	16 (4.4)	39 (9.6)		
>4 years	11 (3.0)	39 (9.6)		

Table 18. Factors related to breast cancer between groups on univariate analysis.

Factors	B	S.E.	Wald	*P* values	OR	95% CI
BMI	-0.797	0.121	43.136	0.000	0.450	0.355-0.572
Water intake	-0.302	0.087	12.180	0.000	0.739	0.624-0.876
Sodium bicarbonate	0.310	0.082	14.419	0.000	**1.363**	1.162-1.599
Fermented pork	0.432	0.118	13.409	0.000	**1.541**	1.223-1.942
Betel quid	-0.334	0.080	17.589	0.000	0.716	0.613-0.837
Cigarette smoking	0.496	0.088	31.939	0.000	**1.643**	1.383-1.951
Secondhand smoke	-0.828	0.156	28.224	0.000	0.437	0.322-0.593
Physical activity	0.307	0.112	7.469	0.006	**1.359**	1.091-1.694
Sleep duration	-0.702	0.125	31.524	0.000	0.496	0.388-0.633
Sleep pattern	0.692	0.157	19.523	0.000	**1.999**	1.470-2.717
Diabetes mellitus	1.428	0.337	17.984	0.000	**4.170**	2.155-8.068
Hypertension	3.217	0.597	29.024	0.000	**24.961**	7.743-80.461
Age at menarche	0.341	0.139	5.982	0.014	**1.406**	1.070-1.848
Menstrual status	1.079	0.156	47.713	0.000	**2.941**	2.166-3.995
Age at menopause	1.054	0.103	103.795	0.000	**2.869**	2.342-3.513
Menstrual duration	-1.036	0.097	113.958	0.000	0.355	0.293-0.429
Dysmenorrhea	-1.093	0.107	103.876	0.000	0.335	0.272-0.414
Pain relief meds	-0.989	0.188	27.695	0.000	0.372	0.258-0.538
Number of births	0.245	0.071	11.812	0.001	**1.278**	1.111-1.470
Breastfeeding	0.456	0.175	6.764	0.009	**1.578**	1.119-2.224
Age at FCB	0.369	0.086	18.376	0.000	**1.446**	1.222-1.712
Age at LCB	0.254	0.073	12.130	0.000	**1.289**	1.117-1.487
Oral contraceptive	-1.224	0.200	37.499	0.000	0.294	0.199-0.435
OC duration	-0.540	0.098	30.521	0.000	0.583	0.481-0.706

5.4. Factors Related to Breast Cancer on Multivariate Analysis

Multivariate logistic regression analysis was done between case and control groups on variables with *P* values less than 0.05 on Pearson's chi-square test and univariate logistic regression. Five factors were significantly associated with breast cancer risk, for which ORs and 95% CIs were fermented pork 2.228 (95% CI: 1.514-3.280, p = 0.000); cigarette smoking 1.528 (95% CI: 1.149-2.033, p = 0.004); hypertension 16.392 (95% CI: 3.296-81.535, p = 0.001); age at menopause 9.809 (95% CI: 4.415-21.789, p = 0.000); age at first child birth 2.732 (95% CI: 1.488-5.018, p = 0.001). But factors like BMI, secondhand smoke, sleep duration, lifetime menstrual

duration and dysmenorrhea seems to have an inverse effect among this study group (Table 19).

Table 19. Factors related to breast cancer between groups on multivariate analysis.

Factors	B	S.E.	Wald	*P* values	OR	95% CI
BMI	-0.991	0.189	27.558	0.000	0.371	0.257-0.538
Water intake	-0.157	0.134	1.365	0.243	0.855	0.657-1.112
Sodium bicarbonate	0.135	0.134	1.018	0.313	1.145	0.880-1.489
Fermented pork	0.801	0.197	16.497	0.000	**2.228**	1.514-3.280
Betel quid	-0.136	0.131	1.076	0.300	0.873	0.675-1.129
Cigarette smoking	0.424	0.146	8.477	0.004	**1.528**	1.149-2.033
Secondhand smoke	-1.055	0.262	16.197	0.000	0.348	0.208-0.582
Physical activity	-0.074	0.189	0.152	0.697	0.929	0.642-1.345
Sleep duration	-0.563	0.207	7.385	0.007	0.570	0.380-0.855
Sleep pattern	0.103	0.282	0.134	0.714	1.109	0.638-1.929
Diabetes mellitus	0.260	0.738	0.124	0.725	1.296	0.305-5.508
Hypertension	2.797	0.818	11.676	0.001	**16.392**	3.296-81.535
Age at menarche	-0.050	0.233	0.046	0.830	0.951	0.602-1.502
Menstrual status	-5.763	1.213	22.585	0.000	0.003	0.000-0.034
Age at menopause	2.283	0.407	31.437	0.000	**9.809**	4.415-21.789
Menstrual duration	-1.845	0.447	17.010	0.000	0.158	0.066-0.380
Dysmenorrhea	-1.370	0.205	44.629	0.000	0.254	0.170-0.380
Pain relief meds	0.515	0.335	2.354	0.125	1.673	0.867-3.228
Number of births	-0.202	0.212	0.903	0.342	0.817	0.539-1.239
Breastfeeding	-0.556	0.895	0.386	0.535	0.574	0.099-3.315
Age at FCB	1.005	0.310	10.506	0.001	**2.732**	1.488-5.018
Age at LCB	-0.116	0.314	0.137	0.712	0.890	0.481-1.647
Oral contraceptive	0.149	0.714	0.044	0.834	1.161	0.287-4.701
OC duration	-0.194	0.322	0.364	0.546	0.824	0.438-1.547

5.5. Comparison of Breast Cancer Cases Based on Survival Status

Survivors are cases with a survival of five years or more after diagnosis and 'deceased' are those who did not survive 5 years or less after diagnosis. Comparison between survivors and deceased was made to check for any significant difference. Significant difference was seen in personal history of cancer wherein the deceased had cancer history of cervix, stomach, breast and brain tumor. Difference was also seen in their first detection and the time taken to consult a doctor after first detection. There

was difference in their level of awareness about breast cancer. Significant association were also seen in lifetime menstrual cycle, cigarette smoking and physical activity (Table 20).

Table 20. Comparison of breast cancer cases on survival status.

Variables	Survived	Deceased	χ^2 (df)	*P* values
Mol subtype			4.561 (4)	0.335
Luminal A	160 (58.6)	51 (68.0)		
Luminal B	35 (12.8)	8 (10.7)		
HER2	31 (11.4)	9 (12.0)		
TrpN	47 (17.2)	7 (9.3)		
Tumor type			6.327 (4)	0.176
IDC	259 (94.9)	73 (97.3)		
ILC	2 (0.7)	2 (2.7)		
DCIS	11 (4.0)	0 (0.0)		
LCIS	1 (0.4)	0 (0.0)		
H/o Cancer			14.308 (1)	0.000
No	283 (100.0)	76 (95.0)		
Yes	0 (0.0)	4 (5.0)		
Cancer in F1			0.205 (2)	0.902
None	155 (54.8)	45 (56.3)		
1 member	87 (30.7)	25 (31.3)		
2 or more	41 (14.5)	10 (12.5)		
Cancer in F2			3.357 (2)	0.187
None	183 (64.7)	60 (75.0)		
1 member	52 (18.4)	12 (15.0)		
2 or more	48 (17.0)	8 (10.0)		
Cancer in F3			3.361 (2)	0.186
None	241 (85.2)	74 (92.5)		
1 member	28 (9.9)	5 (6.3)		
2 or more	14 (4.9)	1 (1.3)		
Types of cancer			4.085 (2)	0.130
None	80 (28.3)	31 (38.8)		
Breast and other	46 (16.3)	8 (10.0)		
Other	157 (55.5)	41 (51.3)		
Symptoms			3.965 (4)	0.411
Lump	265 (93.6)	74 (92.5)		
Nipple Discharge	5 (1.8)	2 (2.5)		
Retracted Nipple	7 (2.5)	2 (2.5)		
Skin Change	6 (2.1)	1 (1.3)		
Ulceration	0 (0.0)	1 (1.3)		

Variables	Survived	Deceased	χ^2 (df)	*P* values
Detection			19.445 (2)	0.000
Self-examination	261 (92.2)	64 (80.0)		
Visible	21 (7.4)	10 (12.5)		
Incidental	1 (0.4)	6 (7.5)		
Reason			4.793 (1)	0.029
Accidental	254 (89.8)	78 (97.5)		
Aware	29 (10.2)	2 (2.5)		
Time to consult			25.556 (2)	0.000
Immediately	214 (75.6)	43 (53.8)		
7-24 months	48 (17.0)	15 (18.8)		
2 years or more	21 (7.4)	22 (27.5)		
Lifetime menstrual cycle			10.015 (2)	0.007
Premenopause	82 (29.0)	34 (42.5)		
<35	125 (44.2)	20 (25.0)		
35 or more	76 (26.9)	26 (32.5)		
Cigarette smoking			15.768 (3)	0.001
No	191 (67.5)	38 (47.5)		
Occassional	30 (10.6)	7 (8.8)		
<10/day	46 (16.3)	24 (30.0)		
>10/day	16 (5.7)	11 (13.8)		
Physical activity			8.720 (2)	0.013
Light	63 (22.3)	12 (15.0)		
Moderate	156 (55.1)	37 (46.3)		
Heavy	64 (22.6)	31 (38.8)		

5.6. Factors Related to Breast Cancer on Cox-regression Model

Survival analysis of all the breast cancer cases was conducted using Cox-regression model to analyze risk factors for variables with *P* values less than 0.05 on Pearson's chi-square test and univariate analysis. Survival of 5 years or more after first detection was assumed as having survived. In this study, three factors were significantly relevant (Table 21) for which ORs and 95% CIs were cigarette smoking 1.272 (95% CI: 1.085-1.491, $p = 0.003$); physical activity 1.499 (95% CI: 1.051-2.139, $p = 0.025$); lifetime cumulative number of menstrual cycles 1.761 (95% CI: 1.246-2.488, $p = 0.001$).

Table 21. Factors related to breast cancer on Cox-regression model.

Factors	B	S.E.	Wald	*P* value	OR	95% CI
BMI	-0.080	0.184	0.187	0.665	0.923	0.644-1.324
Sodium bicarbonate	-0.230	0.144	2.539	0.111	0.795	0.599-1.054
Fermented pork	0.201	0.169	1.407	0.236	1.222	0.877-1.703
Betel quid	-0.032	0.089	0.132	0.716	0.968	0.813-1.153
Cigarette smoking	0.241	0.081	8.811	0.003	**1.272**	1.085-1.491
Secondhand smoke	0.017	0.236	0.005	0.943	1.017	0.640-1.616
Physical activity	0.405	0.181	4.992	0.025	**1.499**	1.051-2.139
Sleep duration	0.288	0.225	1.643	0.200	1.334	0.859-2.073
Sleep pattern	-0.110	0.147	0.562	0.454	0.896	0.671-1.195
Diabetes mellitus	0.263	0.424	0.384	0.535	1.300	0.567-2.983
Hypertension	0.036	0.366	0.010	0.921	1.037	0.506-2.127
Age at menopause	0.150	0.157	0.912	0.340	1.162	0.854-1.581
Lifetime menstrual cycle	0.566	0.177	10.267	0.001	**1.761**	1.246-2.488
Parity	-0.021	0.189	0.012	0.913	0.979	0.676-1.419
Age at FCB	-0.289	0.257	1.263	0.261	0.749	0.453-1.239
Age at LCB	0.276	0.286	0.934	0.334	1.318	0.753-2.307

5.7. Comparison of Groups Based on Menopausal Status

Premenopause

Among premenopausal groups of the study, significant difference was seen in exposure to secondhand smoke, consumption of fruits, physical activity, sleep duration and pattern. There were no significant differences in other variables (Table 22).

Table 22. Comparison of selected characteristics of premenopausal groups.

Premenopausal				
Factors	Ca **n** (%)	Co **n** (%)	χ^2 (df)	*P* values
BMI			0.938 (1)	0.333
≤22.49	105 (56.5)	159 (52.0)		
>22.49	81 (43.5)	147 (48.0)		
Betel quid			3.428 (1)	0.064
Never	33 (17.7)	36 (11.8)		
Ever	153 (82.3)	270 (88.2)		
Tobacco			1.876 (1)	0.171
Never	58 (31.2)	114 (37.3)		
Ever	128 (68.8)	192 (62.7)		
Smoking			3.026 (1)	0.082
Never	133 (71.5)	240 (78.4)		
Ever	53 (28.5)	66 (21.6)		

Factors	Ca **n** (%)	Co **n** (%)	χ^2 (df)	*P* values
Passive smoking			20.752 (1)	0.000
No	78 (41.9)	69 (22.5)		
Yes	108 (58.1)	237 (77.5)		
Tuibur			0.000 (1)	0.983
Never	126 (67.7)	207 (67.6)		
Ever	60 (32.3)	99 (32.4)		
Alcohol			0.842 (1)	0.359
Never	169 (90.9)	285 (93.1)		
Ever	17 (9.1)	21 (6.9)		
Water intake			0.176 (1)	0.675
Low	122 (65.6)	195 (63.7)		
High	64 (34.4)	111 (36.3)		
Salt intake			1.350 (1)	0.245
Low	121 (65.1)	183 (59.8)		
High	65 (34.9)	123 (40.2)		
Fruit			19.109 (1)	0.000
Low	142 (76.3)	174 (56.9)		
High	44 (23.7)	132 (43.1)		
Vegetables			0.499 (1)	0.480
Low	18 (9.7)	24 (7.8)		
High	168 (90.3)	282 (92.2)		
Red meat			1.129 (1)	0.288
Low	182 (97.8)	303 (99.0)		
High	4 (2.2)	3 (1.0)		
Poultry			1.182 (1)	0.277
Low	177 (95.2)	297 (97.1)		
High	9 (4.8)	9 (2.9)		
Seafood			0.858 (1)	0.354
Low	183 (98.4)	297 (97.1)		
High	3 (1.6)	9 (2.9)		
Pork			3.198 (1)	0.074
Low	160 (86.0)	279 (91.2)		
High	26 (14.0)	27 (8.8)		
Fermented pork			1.031 (1)	0.310
Low	175 (94.1)	294 (96.1)		
High	11 (5.9)	12 (3.9)		
Sodium bicarbonate			0.440 (1)	0.507
Low	80 (43.0)	141 (46.1)		
High	106 (57.0)	165 (53.9)		
Physical activity			40.566 (1)	0.000
Normal	125 (67.2)	276 (90.2)		
Heavy	61 (32.8)	30 (9.8)		
Sleep duration			24.964 (1)	0.000
<7 hrs	68 (36.6)	51 (16.7)		
>8 hrs	118 (63.4)	255 (83.3)		
Sleep pattern			11.210 (1)	0.001
Undisturbed	116 (62.4)	234 (76.5)		
Disturbed	70 (37.6)	72 (23.5)		

Association of selected characteristics to breast cancer risk among premenopausal groups

Table 23. Univariate analysis of selected factors on premenopausal groups.

Factors	B	S.E.	Wald	*P* values	OR	95% CI
BMI	-0.181	0.187	0.937	0.333	0.834	0.578-1.204
Betel quid	-0.481	0.261	3.386	0.066	0.618	0.370-1.032
Tobacco	0.270	0.198	1.872	0.171	1.310	0.890-1.930
Cigarette smoking	0.371	0.214	3.010	0.083	1.449	0.953-2.203
Secondhand smoke	-0.909	0.202	20.235	0.000	0.403	0.271-0.599
Tuibur	-0.004	0.199	0.000	0.983	0.996	0.674-1.470
Alcohol	0.311	0.340	0.836	0.360	1.365	0.701-2.660
Water intake	-0.082	0.195	0.176	0.675	0.922	0.629-1.350
Salt intake	-0.224	0.193	1.348	0.246	0.799	0.548-1.167
Fruit intake	-0.895	0.208	18.604	0.000	0.408	0.272-0.614
Vegetable intake	-0.230	0.327	0.497	0.481	0.794	0.419-1.507
Red meat	0.797	0.769	1.074	0.300	2.220	0.491-10.030
Poultry	0.518	0.481	1.159	0.282	1.678	0.654-4.306
Sea food	-0.614	0.673	0.833	0.361	0.541	0.145-2.024
Pork	0.518	0.292	3.148	0.076	1.679	0.947-2.977
Fermented pork	0.432	0.428	1.017	0.313	1.540	0.665-3.565
Sodium bicarbonate	0.124	0.187	0.440	0.507	1.132	0.784-1.635
Physical activity	1.502	0.248	36.761	0.000	**4.490**	2.763-7.295
Sleep duration	-1.058	0.216	23.976	0.000	0.347	0.227-0.530
Sleep pattern	0.674	0.203	11.047	0.001	1.961	1.318-2.918

In this study, among premenopausal groups significant association was seen in physical activity and sleep pattern. But factors like exposure to secondhand smoke, fruit intake and sleep duration seems to have an inverse effect among this study group (Table 23). Among factors with significant association, physical activity was significantly relevant for which OR and 95% CI was 4.515 (95% CI: 2.676-7.617, p = 0.000).

Postmenopause

Among postmenopausal groups, significant difference was seen in BMI, their chewing of betel quid, tobacco, *tuibur* and cigarette smoking, exposure to secondhand

smoke, water and sodium bicarbonate consumed, sleep duration and sleep pattern. There were no significant differences in other variables (Table 24).

Table 24. Comparison of selected characteristics of postmenopausal groups.

Postmenopausal				
Factors	Ca (%)	Co (%)	χ^2 (df)	*P* values
BMI			9.534 (1)	0.002
≤22.49	73 (41.2)	60 (60.6)		
>22.49	104 (58.8)	39 (39.4)		
Betel quid			7.560 (1)	0.006
Never	72 (40.7)	24 (24.2)		
Ever	105 (59.3)	75 (75.8)		
Tobacco			10.196 (1)	0.001
Never	83 (46.9)	27 (27.3)		
Ever	94 (53.1)	72 (72.7)		
Smoking			12.711 (1)	0.000
Never	84 (47.5)	69 (69.7)		
Ever	93 (52.5)	30 (30.3)		
Passive smoking			3.696 (1)	0.055
No	80 (45.2)	33 (33.3)		
Yes	97 (54.8)	66 (66.7)		
Tuibur			4.381 (1)	0.036
Never	124 (70.1)	57 (57.6)		
Ever	53 (29.9)	42 (42.4)		
Alcohol			1.289 (1)	0.256
Never	175 (98.9)	96 (97.0)		
Ever	2 (1.1)	3 (3.0)		
Water intake			9.554 (1)	0.002
Low	129 (72.9)	54 (54.5)		
High	48 (27.1)	45 (45.5)		
Salt intake			0.053 (1)	0.818
Low	121 (68.4)	69 (69.7)		
High	56 (31.6)	30 (30.3)		
Fruit			0.339 (1)	0.561
Low	134 (75.7)	78 (78.8)		
High	43 (24.3)	21 (21.2)		
Vegetables			0.264 (1)	0.607
Low	13 (7.3)	9 (9.1)		
High	164 (92.7)	90 (90.9)		
Red meat			0.561 (1)	0.454
Low	176 (99.4)	99 (100.0)		
High	1 (0.6)	0 (0.0)		
Poultry			4.017 (1)	0.045
Low	170 (96.0)	99 (100.0)		
High	7 (4.0)	0 (0.0)		

Factors	Ca (%)	Co (%)	χ^2 (df)	*P* values
Seafood			0.155 (1)	0.693
Low	170 (96.0)	96 (97.0)		
High	7 (4.0)	3 (3.0)		
Pork			0.030 (1)	0.862
Low	162 (91.5)	90 (90.9)		
High	15 (8.5)	9 (9.1)		
Fermented pork			3.087 (1)	0.079
Low	162 (91.5)	96 (97.0)		
High	15 (8.5)	3 (3.0)		
Sodium bicarbonate			9.338 (1)	0.002
Low	63 (35.6)	54 (54.5)		
High	114 (64.4)	45 (45.5)		
Physical activity			2.244 (1)	0.134
Normal	123 (69.5)	60 (60.6)		
Heavy	54 (30.5)	39 (39.4)		
Sleep duration			11.873 (1)	0.001
<7 hrs	80 (45.2)	24 (24.2)		
>8 hrs	97 (54.8)	75 (75.8)		
Sleep pattern			4.285 (1)	0.038
Undisturbed	101 (57.1)	69 (69.7)		
Disturbed	76 (42.9)	30 (30.3)		

Association of selected characteristics to breast cancer risk among postmenopausal groups

Among postmenopausal groups, significant association was seen in BMI, cigarette smoking, sodium bicarbonate and sleep pattern. But factors like chewing of betel quid, tobacco and *tuibur*, water intake and sleep duration seem to have an inverse effect among this study group (Table 25). Among factors with significant association, three factors were significantly relevant for which ORs and 95% CIs were BMI 2.145 (95% CI: 1.205-3.817, p = 0.009), cigarette smoking 2.968 (95% CI: 1.613-5.460, p = 0.000) and consumption of sodium bicarbonate 2.457 (95% CI: 1.381-4.371, p = 0.002).

Table 25. Univariate analysis of selected factors on postmenopausal groups.

Factors	B	S.E.	Wald	*P* values	OR	95% CI
BMI	0.785	0.256	9.384	0.002	**2.192**	1.327-3.621
Betel quid	-0.762	0.280	7.408	0.006	0.467	0.270-0.808
Tobacco	-0.856	0.271	9.963	0.002	0.425	0.250-0.723
Cigarette smoking	0.935	0.265	12.395	0.000	**2.546**	1.513-4.285
Passive smoking	-0.500	0.261	3.669	0.055	0.606	0.363-1.012
Tuibur	-0.545	0.261	4.343	0.037	0.580	0.348-0.968
Alcohol	-1.006	0.922	1.191	0.275	0.366	0.060-2.227
Water intake	-0.806	0.263	9.377	0.002	0.447	0.267-0.748
Salt intake	0.062	0.272	0.053	0.818	1.064	0.625-1.814
Fruit intake	0.176	0.302	0.338	0.561	1.192	0.660-2.154
Vegetable intake	0.232	0.453	0.263	0.608	1.262	0.519-3.066
Sea food	0.276	0.702	0.155	0.694	1.318	0.333-5.214
Pork	-0.077	0.442	0.030	0.862	0.926	0.390-2.200
Fermented pork	1.086	0.645	2.832	0.092	2.963	0.836-10.498
Sodium bicarbonate	0.775	0.256	9.195	0.002	**2.171**	1.315-3.584
Physical activity	-0.392	0.263	2.233	0.135	0.675	0.404-1.130
Sleep duration	-0.947	0.279	11.520	0.001	0.388	0.225-0.670
Sleep pattern	0.549	0.266	4.245	0.039	**1.731**	1.027-2.916

5.8. History of Breast and Other Cancers

Among the study group, there were 54 cases with family history of breast cancer and 309 sporadic cases with no history of breast cancer cases even up to third degree relatives. Among 309 sporadic cases, 111 were with no history of any type of cancer in their family even to third degree and 198 were with other types of cancer. In this study, 14.9 percent of cases had family history of breast cancer, 54.5 percent with family history of other cancer, which means 69.4 percent were with history of breast or other cancer. 30.6 percent of cases were with no history of breast or other cancer even up to third degree relatives (Table 26).

Table 26. Cases with number of I, II- and III-degree relatives having cancer.

Type of Cancer	F1	F2	F3
Breast	12	10	16
Breast and other cancer	7	8	4
Other cancer	140	97	24
Suspect	4	5	4
None	200	243	315

5.9. Molecular Study

5.9.1. Comparison of Cases and Controls Selected for Molecular Analysis

Among the study groups, significant differences were seen in hypertension, pesticide exposure, sleep duration and pattern, age at menopause, lifetime cumulative number of menstrual cycles, breastfeeding duration, age at first child birth. There were no other significant differences in other factors (Table 27, 28 and 29).

Table 27. Comparison of selected variables between groups for molecular analysis.

Variables	Case **n** (%)	Control **n** (%)	$\chi2$ (df)	*P* value
Residence			0.569 (1)	0.450
Urban	26 (53.1)	25 (61.0)		
Rural	23 (46.9)	16 (39.0)		
Education			3.138 (1)	0.076
<9	27 (55.1)	30 (73.2)		
>10	22 (44.9)	11 (26.8)		
Marital status			1.689 (1)	0.194
Single	8 (16.3)	3 (7.3)		
Married	41 (83.7)	38 (92.7)		
BMI			0.105 (1)	0.746
<21	14 (28.6)	13 (31.7)		
>21.1	35 (71.4)	28 (68.3)		
Diabetes mellitus			1.496 (1)	0.221
No	43 (87.8)	39 (95.1)		
Yes	6 (12.2)	2 (4.9)		
Hypertension			6.351 (1)	0.012
No	42 (85.7)	41 (100)		
Yes	7 (14.3)	0 (0.0)		
Asthma & Allergy			0.846 (1)	0.358
No	48 (98.0)	41 (100)		
Yes	1 (2.0)	0 (0.0)		
Pesticide exposure			4.186 (1)	0.041
Unexposed	47 (95.9)	34 (82.9)		
Exposed	2 (4.1)	7 (17.1)		
EM exposure			0.403 (1)	0.525
Unexposed	26 (53.1)	19 (46.3)		
Exposed	23 (46.9)	22 (53.7)		

Table 28. Comparison of dietary and behavioral characteristics for molecular analysis.

Variables	Case **n** (%)	Control **n** (%)	$\chi2$ (df)	*P* value
Water intake			1.698 (1)	0.193
Low	34 (69.4)	23 (56.1)		
High	15 (30.6)	18 (43.9)		
Salt intake			0 (1)	0.988
Low	31 (63.3)	26 (63.4)		
High	18 (36.7)	15 (36.6)		
Sodium bicarbonate			0.91 (1)	0.340
Low	19 (38.8)	20 (48.8)		
High	30 (61.2)	21 (51.2)		
Fruits			3.009 (1)	0.083
Low	42 (85.7)	29 (70.7)		
High	7 (14.3)	12 (29.3)		
Vegetables			0.07 (1)	0.791
Low	4 (8.2)	4 (9.8)		
High	45 (91.8)	37 (90.2)		
Red meat			0.846 (1)	0.358
Low	48 (98.0)	41 (100)		
High	1 (2.0)	0 (0.0)		
Poultry			0.016 (1)	0.898
Low	48 (98.0)	40 (97.6)		
High	1 (2.0)	1 (2.4)		
Seafood			0.016 (1)	0.898
Low	48 (98.0)	40 (97.6)		
High	1 (2.0)	1 (2.4)		
Pork			0.427 (1)	0.514
Low	42 (85.7)	37 (90.2)		
High	7 (14.3)	4 (9.8)		
Fermented pork			2.962 (1)	0.085
Low	41 (83.7)	39 (95.1)		
High	8 (16.3)	2 (4.9)		
Betel quid			0.890 (1)	0.346
Never	11 (22.4)	6 (14.6)		
Ever	38 (77.6)	35 (85.4)		
Tobacco			2.920 (1)	0.087
Never	11 (22.4)	16 (39.0)		
Ever	38 (77.6)	25 (61.0)		

Variables	Case **n** (%)	Control **n** (%)	χ2 (df)	*P* value
Smoke			1.934 (1)	0.164
Never	29 (59.2)	30 (73.2)		
Ever	20 (40.8)	11 (26.8)		
Tuibur			0.894 (1)	0.345
Never	30 (61.2)	29 (70.7)		
Ever	19 (38.8)	12 (29.3)		
Alcohol			2.163 (1)	0.141
Never	44 (89.8)	40 (97.6)		
Ever	5 (10.2)	1 (2.4)		
Secondhand smoke			1.775 (1)	0.183
No	21 (42.9)	12 (29.3)		
Yes	28 (57.1)	29 (70.7)		
Physical activity			0.179 (1)	0.672
Normal	40 (81.6)	32 (78.0)		
Heavy	9 (18.4)	9 (22.0)		
Sleep duration			9.115 (1)	0.003
<7 hrs	20 (40.8)	5 (12.2)		
>8 hrs	29 (59.2)	36 (87.8)		
Sleep pattern			4.363 (1)	0.037
Undisturbed	24 (49.0)	29 (70.7)		
Disturbed	25 (51.0)	12 (29.3)		

5.9.2. Association of Selected Variables to Breast Cancer Risk Among Groups

The groups selected for molecular study was further analyzed to examine the causal effect of some of the factors. Factors were further adjusted for age and residence but the results did not show much differences. Stratification of variables based on a single common referent group of characteristics was also used (Table 30).

Table 29. Comparison of reproductive characteristics for molecular analysis.

Variables	Case **n** (%)	Control **n** (%)	$\chi 2$ (df)	*P* value
Age at menarche			3.529 (2)	0.171
<12	4 (8.2)	0 (0.0)		
13-15 years	30 (61.2)	28 (68.3)		
16>	15 (30.6)	13 (31.7)		
Menstrual status			0.569 (1)	0.450
Premenopausal	26 (53.1)	25 (61.0)		
Postmenopausal	23 (46.9)	16 (39.0)		
Age at menopause			10.851 (3)	0.013
Premenopause	14 (28.6)	25 (61.0)		
<49	21 (42.9)	10 (24.4)		
50-54	11 (22.4)	6 (14.6)		
55>	3 (6.1)	0 (0.0)		
Menstrual duration			9.546 (2)	0.008
Premenopause	22 (44.9)	10 (24.4)		
<35	13 (26.5)	6 (14.6)		
>36	14 (28.6)	25 (61.0)		
Miscarriage			0.206 (1)	0.650
Never	30 (61.2)	27 (65.9)		
Ever	19 (38.8)	14 (34.1)		
Parity			2.359 (1)	0.125
Nulliparous	9 (18.4)	3 (7.3)		
Parous	40 (81.6)	38 (92.7)		
Breastfeeding			2.359 (1)	0.125
Never	9 (18.4)	3 (7.3)		
Ever	40 (81.6)	38 (92.7)		
BF duration			17.662 (3)	0.001
0	9 (18.4)	3 (7.3)		
≤6 months	0 (0.0)	1 (2.4)		
7-12 months	15 (30.6)	1 (2.4)		
≥13 months	25 (51.0)	36 (87.8)		
Age at FCB			9.603 (3)	0.022
0	9 (18.4)	3 (7.3)		
≤19	4 (8.2)	12 (29.3)		
20-34	34 (69.4)	22 (53.7)		
≥35	2 (4.1)	4 (9.8)		
Age at LCB			5.021 (3)	0.170
0	9 (18.4)	3 (7.3)		
≤19	1 (2.0)	0 (0.0)		
20-34	29 (59.2)	23 (56.1)		
≥35	10 (20.4)	15 (36.6)		
Contraceptive use			1.643 (1)	0.200
Never	35 (71.4)	24 (58.5)		
Ever	14 (28.6)	17 (41.5)		

Table 30. Selected variables of the study subjects with odd ratio and adjusted OR.

Variables	Case **n** (%)	Control **n** (%)	Odds Ratio (95% CI)	Adjusted OR (95% CI)
Parity				
Nulliparous	9 (18.4)	3 (7.3)	1.00	1.00
1-2 children	10 (20.4)	8 (19.5)	3.33 (1.33-8.33)	6.01 (2.17-16.66)
3-4 children	21 (42.9)	20 (48.8)	1.39 (0.66-2.93)	2.33 (0.99-5.51)
5 or more	9 (18.4)	10 (24.4)	1.17 (0.62-2.19)	1.52 (0.77-3.01)
Breastfeeding				
Never	9 (18.4)	3 (7.3)	1.00	1.00
Ever	40 (81.6)	38 (92.7)	2.85 (1.29-6.32)	3.86 (1.63-9.14)
Age at FCB				
Nulliparous	9 (18.4)	3 (7.3)	1.00	1.00
≤19	4 (8.2)	12 (29.3)	6.00 (1.74-20.67)	7.36 (2.00-27.09)
20-34	34 (69.4)	22 (53.7)	0.67 (0.21-2.16)	0.70 (0.21-2.33)
≥35	2 (4.1)	4 (9.8)	3.09 (1.11-8.64)	3.00 (1.04-8.60)
Contraceptive use				
Never	35 (71.4)	24 (58.5)	1.00	1.00
Ever	14 (28.6)	17 (41.5)	1.77 (1.07-2.94)	1.52 (0.89-2.61)
Contraceptive duration				
Never	35 (71.4)	24 (58.5)	1.00	1.00
<1 year	3 (6.1)	1 (2.4)	6.56 (2.58-16.70)	5.64 (2.17-14.62)
1-4 years	9 (18.4)	7 (17.1)	13.50 (2.79-65.41)	13.50 (2.78-65.58)
>4 years	2 (4.1)	9 (22.0)	5.79 (2.02-16.57)	5.79 (2.02-16.61)
BMI				
≤22.49	21 (42.9)	20 (48.8)	1.00	1.00
>22.49	28 (57.1)	21 (51.2)	0.79 (0.49-1.27)	0.79 (0.48-1.29)
Betel quid				
Never	11 (22.4)	6 (14.6)	1.00	1.00
Ever	38 (77.6)	35 (85.4)	1.69 (0.90-3.18)	1.54 (0.79-3.00)
Tobacco				
Never	11 (22.4)	16 (39.0)	1.00	1.00
Ever	38 (77.6)	25 (61.0)	0.45 (0.27-0.77)	0.41 (0.24-0.71)
Smoke				
Never	29 (59.2)	30 (73.2)	1.00	1.00
Ever	20 (40.8)	11 (26.8)	0.53 (0.32-0.89)	0.59 (0.35-1.01)
Tuibur				
Never	30 (61.2)	29 (70.7)	1.00	1.00
Ever	19 (38.8)	12 (29.3)	0.65 (0.39-1.09)	0.70 (0.41-1.17)

Variables	Case **n** (%)	Control **n** (%)	Odds Ratio (95% CI)	Adjusted OR (95% CI)
Alcohol				
Never	44 (89.8)	40 (97.6)	1.00	1.00
Ever	5 (10.2)	1 (2.4)	0.22 (0.06-0.78)	0.17 (0.05-0.61)
Secondhand smoke				
No	21 (42.9)	12 (29.3)	1.00	1.00
Yes	28 (57.1)	29 (70.7)	1.81 (1.09-3.01)	1.90 (1.13-3.17)
Water intake				
High	15 (30.6)	18 (43.9)	1.00	1.00
Low	34 (69.4)	23 (56.1)	1.77 (1.08-2.92)	1.96 (1.17-3.27)
Fruits intake				
High	7 (14.3)	12 (29.3)	1.00	1.00
Low	42 (85.7)	29 (70.7)	2.48 (1.36-4.54)	2.27 (1.23-4.19)
Vegetables intake				
High	45 (91.8)	37 (90.2)	1.00	1.00
Low	4 (8.2)	4 (9.8)	0.82 (0.36-1.90)	0.79 (0.34-1.85)
Salt intake				
Low	31 (63.3)	26 (63.4)	1.00	1.00
High	18 (36.7)	15 (36.6)	0.99 (0.60-1.63)	0.95 (0.57-1.57)
Fermented pork				
Low	41 (83.7)	39 (95.1)	1.00	1.00
High	8 (16.3)	2 (4.9)	0.26 (0.10-0.67)	0.22 (0.08-0.57)
Sodium bicarbonate				
Low	19 (38.8)	20 (48.8)	1.00	1.00
High	30 (61.2)	21 (51.2)	0.67 (0.41-1.08)	0.60 (0.36-0.98)
Sleep duration				
≥8 hours	29 (59.2)	36 (87.8)	1.00	1.00
≤7 hours	20 (40.8)	5 (12.2)	4.97 (2.64-9.35)	5.15 (2.70-9.81)
Sleep pattern				
Undisturbed	24 (49.0)	29 (70.7)	1.00	1.00
Disturbed	25 (51.0)	12 (29.3)	0.40 (0.24-0.66)	0.41 (0.24-0.68)
Menstrual duration				
Premenopause	22 (44.9)	10 (24.4)	1.00	1.00
<35	13 (26.5)	6 (14.6)	3.93 (2.21-6.97)	6.97 (3.26-14.87)
>36	14 (28.6)	25 (61.0)	3.87 (1.97-7.59)	8.95 (3.49-22.95)

5.9.3. Glutathione S-transferase

The presence of fragments at 480 bp for GSTT1 and 215 bp for GSTM1 (Figure 9) identified the subject as having wild genotype. Whereas, the absence of fragments at 480 bp for GSTT1 and 215 bp for GSTM1 (Figure 9) identified the subject as having null genotype.

In this study, GSTT1 null genotype was 40.8 percent and 68.3 percent in case and control groups respectively, overall prevalence of null genotype in the study group was 53.3 percent (Table 31). There was significant difference between case and control groups. However, there were more cases with GSTT1 wild genotype than null genotype, and a greater number of controls with null genotype than wild type. When comparison was made between case and control groups with null genotype, difference was significant in frequency of consumption of fermented pork and hypertension; with wild genotype, sleep duration and cigarette smoking.

In this study, GSTM1 null genotype was 63.3 percent and 53.7 percent in case and control groups respectively, overall prevalence of null genotype in the study group was 58.9 percent (Table 31). Comparison between the groups did not show any significant differences. When comparison was made between case and control groups with null genotype, difference was significant in sleep duration, pesticide exposure and hypertension; with wild genotype, sleep duration, sleep pattern and cigarette smoking.

The presence or absence in combination of GSTT1 and GSTM1 shows significant differences. There were a greater number of controls with homozygous deletions in this study group. Overall homozygous deletions of the subjects were 31.1 percent, 26.5 percent of cases and 36.6 percent of controls and could not be associated with increased breast cancer risk (Table 31).

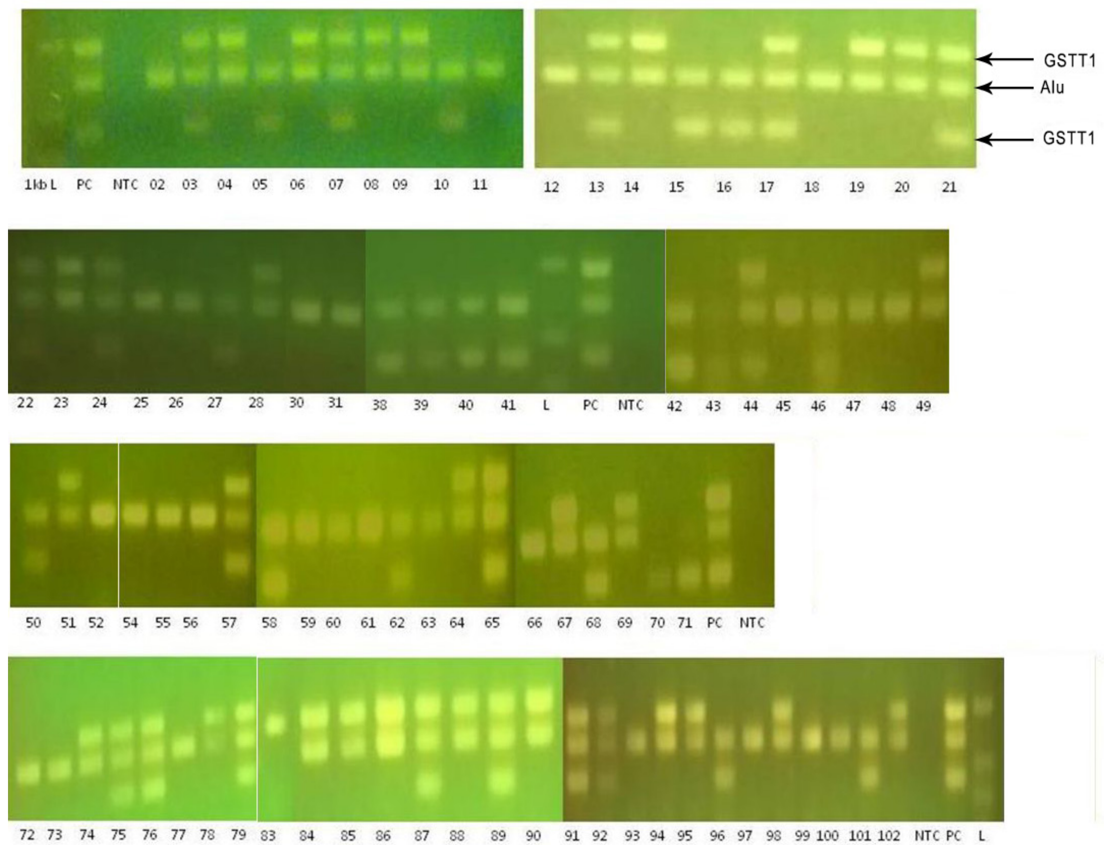


Figure 9. Results of multiplex PCR for GSTM1 and GSTT1.

[Case = 2-28, 30-31, 83-102; Control = 38-52, 54-79;
NTC = No template control; PC = Positive control; L = Ladder]

Table 31. Comparison of GSTT1 and GSTM1 between groups.

Gene	Case **n** (%)	Control **n** (%)	$\chi 2$ (df)	*P* values
GSTT1			20.313 (1)	0.000
Normal	29 (59.2)	13 (31.7)		
Null	20 (40.8)	28 (68.3)		
GSTM1			2.553 (1)	0.110
Normal	18 (36.7)	19 (46.3)		
Null	31 (63.3)	22 (53.7)		
T1 & M1			22.807 (3)	0.000
T&M	11 (22.5)	6 (14.6)		
T&Null	18 (36.7)	7 (17.1)		
Null&M	7 (14.3)	13 (31.7)		
Null&Null	13 (26.5)	15 (36.6)		

Association of selected variables with GSTT1 and GSTM1 genotype

GSTT1

As mentioned before, in spite of significant difference among groups, a greater number of controls were with GSTT1 null genotype. Interaction of some of the risk factors with GSTT1 genotype did not find any significant association. The findings from this statistical analysis could not conclude that breast cancer risk increases with having null genotype. In this study, no definite conclusion could be made that certain established risk factors with having null genotype influence increased risk (Table 34). Even when stratified on their menopausal status (Table 32), there was no association of null genotype with breast cancer risk as concluded in other studies.

GSTM1

As mentioned above, there was no significant difference among the study groups on GSTM1 genotype. Interaction of some of the risk factors with GSTM1 genotype did not find any significant association. The findings from this statistical analysis could not conclude that breast cancer risk increases with having null genotype. In this study, no definite conclusion could be made that certain established risk factors with having null genotype influence increased risk, except sleep duration of less than 7 hours seems to influence risk (Table 35). Even when stratified on their menopausal status (Table 32), there was no association of null genotype with breast cancer risk as concluded in other studies.

Table 32. Association of GSTT1 and GSTM1 genotype on menopausal status.

	GSTT1			GSTM1	
	Normal	Null		Normal	Null
Total					
Case	29 (59.2)	20 (40.8)		18 (36.7)	31 (63.3)
Control	13 (31.7)	28 (68.3)		19 (46.3)	22 (53.7)
OR[1] (95% CI)	1.00	0.31 (0.18-0.53)		1.00	1.71 (1.03-2.84)
$\chi 2$ *P*	0.000			0.110	
Premenopausal					
Case	15 (57.7)	11 (42.3)		9 (34.6)	17 (65.4)
Control	9 (36.0)	16 (64.0)		11 (44.0)	14 (56.0)
OR[1] (95% CI)	1.00	0.29 (0.13-0.66)		1.00	1.55 (0.79-3.05)
$\chi 2$ *P*	0.007			0.235	
Postmenopausal					
Case	14 (60.9)	9 (39.1)		9 (39.1)	14 (60.9)
Control	4 (25.0)	12 (75.0)		8 (50.0)	8 (50.0)
OR[1] (95% CI)	1.00	0.20 (0.08-0.50)		1.00	2.30 (0.99-5.35)
$\chi 2$ *P*	0.000			0.224	

Table 33. Association of GSTT1/GSTM1 genotype on menopausal status.

	M1T1 Normal	T Normal & M Null	T Null & M Normal	M1T1 Null
Total				
Case	11 (22.5)	18 (36.7)	7 (14.3)	13 (26.5)
Control	6 (14.6)	7 (17.1)	13 (31.7)	15 (36.6)
OR[1] (95% CI)	1.00	1.88 (0.88-3.99)	3.17 (1.59-6.30)	0.58 (0.29-1.17)
$\chi 2$ *P*	0.000			
Premenopausal				
Case	5 (19.2)	10 (38.5)	4 (15.4)	7 (26.9)
Control	4 (16.0)	5 (20.0)	7 (28.0)	9 (36.0)
OR[1] (95% CI)	1.00	2.24 (0.75-6.64)	3.39 (1.33-8.65)	0.58 (0.21-1.56)
$\chi 2$ *P*	0.037			
Postmenopausal				
Case	6 (26.1)	8 (34.8)	3 (13.0)	6 (26.1)
Control	2 (12.5)	2 (12.5)	6 (37.5)	6 (37.5)
OR[1] (95% CI)	1.00	2.28 (0.64-8.08)	3.63 (1.11-11.94)	0.31 (0.10-1.00)
$\chi 2$ *P*	0.001			

[1] Subjects with GSTT1 and GSTM1 wild genotype serve as a referent, adjusted for age, age at menarche, BMI, residence

Table 34. Association of GSTT1 with selected variables.

Factors	GSTT1	Ca (%)	Co (%)	OR (95% CI)
BMI				
≤22.49	Normal	16 (32.7)	6 (14.6)	1.00
>22.49	Normal	13 (26.5)	7 (17.1)	1.363 (1.045-1.777)
≤22.49	Null	5 (10.2)	14 (34.1)	0.357 (0.198-0.644)
>22.49	Null	15 (30.6)	14 (34.1)	1.035 (0.839-1.277)
Betel quid				
Never	Normal	8 (16.3)	2 (4.9)	1.00
Ever	Normal	21 (42.9)	11 (26.8)	1.382 (1.119-1.706)
Never	Null	3 (6.1)	4 (9.8)	0.750 (0.316-1.780)
Ever	Null	17 (34.7)	24 (58.5)	0.842 (0.703-1.007)
Tobacco				
Never	Normal	6 (12.2)	4 (9.8)	1.00
Ever	Normal	23 (46.9)	9 (22.0)	1.599 (1.280-1.997)
Never	Null	5 (10.2)	12 (29.3)	0.417 (0.228-0.761)
Ever	Null	15 (30.6)	16 (39.0)	0.968 (0.790-1.187)
Smoking				
Never	Normal	14 (28.6)	11 (26.8)	1.00
Ever	Normal	15 (30.6)	2 (4.9)	2.739 (1.789-4.193)
Never	Null	15 (30.6)	19 (46.3)	0.789 (0.534-1.167)
Ever	Null	5 (10.2)	9 (22.0)	0.745 (0.544-1.022)
Passive smoking				
No	Normal	10 (20.4)	4 (9.8)	1.00
Yes	Normal	19 (38.8)	9 (22.0)	1.453 (1.156-1.827)
No	Null	11 (22.4)	8 (19.5)	1.375 (0.813-2.326)
Yes	Null	9 (18.4)	20 (48.8)	0.671 (0.535-0.842)
Tuibur				
Never	Normal	16 (32.7)	10 (24.4)	1.00
Ever	Normal	13 (26.5)	3 (7.3)	2.082 (1.449-2.991)
Never	Null	14 (28.6)	19 (46.3)	0.737 (0.495-1.098)
Ever	Null	6 (12.2)	9 (22.0)	0.816 (0.606-1.100)
Water intake				
High	Normal	24 (49.0)	7 (17.1)	1.00
Low	Normal	5 (10.2)	6 (14.6)	3.429 (2.109-5.575)
High	Null	10 (20.4)	16 (39.0)	0.913 (0.716-1.163)
Low	Null	10 (20.4)	12 (29.3)	0.625 (0.396-0.986)

Factors	GSTT1	Ca (%)	Co (%)	OR (95% CI)
Salt intake				
Low	Normal	17 (34.7)	7 (17.1)	1.00
High	Normal	12 (24.5)	6 (14.6)	1.414 (1.066-1.877)
Low	Null	14 (28.6)	19 (46.3)	0.737 (0.495-1.098)
High	Null	6 (12.2)	9 (22.0)	0.816 (0.606-1.100)
Fermented pork				
Low	Normal	25 (51.0)	11(24.5)	1.00
High	Normal	4 (8.2)	2(24.5)	1.41 (0.61-3.30)
Low	Null	16 (32.7)	28(24.5)	0.57 (0.31-1.06)
High	Null	4 (8.2)	0(0.0)	1.00
Sleep duration				
≥8 hrs	Normal	15 (30.6)	12 (29.3)	1.00
≤7 hrs	Normal	14 (28.6)	1 (2.4)	14.000 (4.339-45.167)
≥8 hrs	Null	14 (28.6)	24 (58.5)	0.764 (0.631-0.924)
≤7 hrs	Null	6 (12.2)	4 (9.8)	1.500 (0.723-3.114)
Sleep pattern				
Undisturbed	Normal	13 (26.5)	9 (22.0)	1.00
Disturbed	Normal	16 (32.7)	4 (9.8)	2.000 (1.458-2.744)
Undisturbed	Null	11 (22.4)	20 (48.8)	0.550 (0.360-0.841)
Disturbed	Null	9 (18.4)	8 (19.5)	1.061 (0.806-1.396)
Miscarriage				
Never	Normal	14 (28.6)	8 (19.5)	1.00
Ever	Normal	15 (30.6)	5 (12.2)	1.732 (1.293-2.320)
Never	Null	16 (32.7)	19 (46.3)	0.842 (0.574-1.236)
Ever	Null	4 (8.2)	9 (22.0)	0.667 (0.475-0.937)
Pesticide exposure				
Unexposed	Normal	27 (55.1)	10 (24.4)	1.00
Exposed	Normal	2 (4.1)	3 (7.3)	0.82 (0.33-1.20)
Unexposed	Null	20 (40.8)	24 (58.5)	0.83 (0.46-1.51)
Exposed	Null	0(0.0)	4 (9.8)	1.00

Table 35. Association of GSTM1 with selected variables.

Factors	GSTM1	Ca (%)	Co (%)	OR (95% CI)
BMI				
≤22.49	Normal	6 (12.2)	9 (22.0)	1.00
>22.49	Normal	12 (24.5)	10 (24.4)	1.095 (0.860-1.396)
≤22.49	Null	15 (30.6)	11 (26.8)	1.364 (0.870-2.137)
>22.49	Null	16 (32.7)	11 (26.8)	1.206 (0.966-1.505)
Betel quid				
Never	Normal	4 (8.2)	2 (4.9)	1.00
Ever	Normal	14 (28.6)	17 (41.5)	0.907 (0.740-1.113)
Never	Null	7 (14.3)	4 (9.8)	1.750 (0.861-3.557)
Ever	Null	24 (49.0)	18 (43.9)	1.155 (0.968-1.377)
Tobacco				
Never	Normal	2 (4.1)	7 (17.1)	1.00
Ever	Normal	16 (32.7)	12 (29.3)	1.155 (0.930-1.433)
Never	Null	9 (18.4)	9 (22.0)	1.00
Ever	Null	22 (44.9)	13 (31.7)	1.301 (1.067-1.586)
Smoking				
Never	Normal	6 (12.2)	15 (36.6)	1.00
Ever	Normal	12 (24.5)	4 (9.8)	1.732 (1.249-2.401)
Never	Null	23 (46.9)	15 (36.6)	1.533 (1.053-2.232)
Ever	Null	8 (16.3)	7 (17.1)	1.069 (0.798-1.433)
Passive smoking				
No	Normal	8 (16.3)	5 (12.2)	1.00
Yes	Normal	10 (20.4)	14 (34.1)	0.845 (0.669-1.068)
No	Null	13 (26.5)	7 (17.1)	1.857 (1.093-3.157)
Yes	Null	18 (36.7)	15 (36.6)	1.095 (0.899-1.335)
Tuibur				
Never	Normal	10 (20.4)	14 (34.1)	1.00
Ever	Normal	8 (16.3)	5 (12.2)	1.265 (0.916-1.746)
Never	Null	20 (40.8)	15 (36.6)	1.333 (0.906-1.962)
Ever	Null	11 (22.4)	7 (17.1)	1.254 (0.954-1.648)
Water intake				
High	Normal	14 (28.6)	9 (22.0)	1.00
Low	Normal	4 (8.2)	10 (24.4)	1.429 (0.963-2.119)
High	Null	20 (40.8)	14 (34.1)	0.632 (0.453-0.884)
Low	Null	11 (22.4)	8 (19.5)	1.556 (0.959-2.523)

Factors	GSTM1	Ca (%)	Co (%)	OR (95% CI)
Salt intake				
Low	Normal	12 (24.5)	11 (26.8)	1.00
High	Normal	6 (12.2)	8 (19.5)	0.866 (0.638-1.176)
Low	Null	19 (38.8)	15 (36.6)	1.267 (0.857-1.872)
High	Null	12 (24.5)	7 (17.1)	1.309 (1.000-1.714)
Sodium bicarbonate				
Low	Normal	6 (12.2)	11 (26.8)	1.00
High	Normal	12 (24.5)	8 (19.5)	1.225 (0.946-1.586)
Low	Null	13 (26.5)	9 (22.0)	1.444 (0.884-2.359)
High	Null	18 (36.7)	13 (31.7)	1.177 (0.958-1.446)
Fermented pork				
Low	Normal	16 (32.7)	18 (43.9)	1.00
High	Normal	2 (4.1)	1 (2.4)	1.41 (0.43-4.70)
Low	Null	25 (51.0)	21 (51.2)	1.19 (0.67-2.13)
High	Null	6 (12.2)	1 (2.4)	2.45 (0.85-7.06)
Sleep pattern				
Undisturbed	Normal	9 (18.4)	16 (39.0)	1.00
Disturbed	Normal	9 (18.4)	3 (7.3)	1.732 (1.188-2.526)
Undisturbed	Null	15 (30.6)	13 (31.7)	1.154 (0.751-1.772)
Disturbed	Null	16 (32.7)	9 (22.0)	1.333 (1.053-1.688)
Miscarriage				
Never	Normal	10 (20.4)	13 (31.7)	1.00
Ever	Normal	8 (16.3)	6 (14.6)	1.155 (0.851-1.567)
Never	Null	20 (40.8)	14 (34.1)	1.429 (0.963-2.119)
Ever	Null	11 (22.4)	8 (19.5)	1.173 (0.902-1.525)
Pesticide exposure				
Unexposed	Normal	17 (34.7)	17 (41.5)	1.00
Exposed	Normal	1 (2.0)	2 (4.9)	0.72 (0.21-2.35)
Unexposed	Null	30 (61.3)	17 (41.5)	1.77 (0.97-3.20)
Exposed	Null	1 (2.0)	5 (12.2)	0.45 (0.15-1.31)

5.9.4. Mitochondrial DNA

D-loop

There were more variations per sample among case groups than control groups. A total of 114 sequence variations at 109 distinct nucleotide positions were found in the 90 samples, 35 in HVI region and 74 in HVII region. Among these variations, there were 6 deletions (HVI = 2, HVII = 4); 6 insertions (HVI = 3, HVII = 3); 4 transversion (HVI = 1, HVII = 3); 98 transitions (HVI = 30, HVII = 68) as shown in Table 36.

Table 36. Comparison of variations found in D-loop region of case and control.

Locus	Base Change	Ref Seq	Mut Type	Case n (%)	Control n (%)	Locus	Base Change	Ref Seq	Mut Type	Case n (%)	Control n (%)
46	T>G	T	Tv	**1** (2.04)	**0** (0)	**16182**	A>ACC	A	Int	**0** (0)	**1** (2.44)
73	A>G	A	Ts	**49** (100)	**41** (100)	**16183**	A>C	A	Tv	**10** (20.41)	**6** (14.63)
94	G>A	G	Ts	**0** (0)	**2** (4.88)	**16183**	A DEL	A	Del	**1** (2.04)	**0** (0)
143	G>A	G	Ts	**1** (2.04)	**0** (0)	**16185**	C>T	C	Ts	**5** (10.20)	**2** (4.88)
146	T>C	T	Ts	**15** (30.61)	**7** (17.07)	**16187**	C>T	C	Ts	**0** (0)	**1** (2.44)
150	C>T	C	Ts	**2** (4.08)	**4** (9.76)	**16189**	T>C	T	Ts	**12** (24.49)	**6** (14.63)
151	C>T	C	Ts	**5** (10.20)	**6** (14.63)	**16189**	T DEL	T	Del	**3** (6.12)	**2** (4.88)
152	T>C	T	Ts	**20** (40.82)	**14** (34.15)	**16193**	C>CC	C	Int	**4** (8.16)	**0** (0)
153	A>G	A	Ts	**2** (4.08)	**2** (4.88)	**16209**	T>C	T	Ts	**0** (0)	**1** (2.44)
173	T>C	T	Ts	**0** (0)	**1** (2.44)	**16213**	G>A	G	Ts	**3** (6.12)	**0** (0)
183	A>G	A	Ts	**3** (6.12)	**2** (4.88)	**16214**	C>T	C	Ts	**0** (0)	**1** (2.44)
184	G>A	G	Ts	**3** (6.12)	**1** (2.44)	**16217**	T>C	T	Ts	**0** (0)	**3** (7.32)
185	G>A	G	Ts	**7** (14.29)	**5** (12.20)	**16218**	C>T	C	Ts	**4** (8.16)	**1** (2.44)
189	A>G	A	Ts	**0** (0)	**1** (2.44)	**16223**	C>T	C	Ts	**25** (51.02)	**28** (68.29)
195	T>C	T	Ts	**4** (8.16)	**4** (9.76)	**16224**	T>C	T	Ts	**0** (0)	**1** (2.44)
199	T>C	T	Ts	**1** (2.04)	**0** (0)	**16225**	C>T	C	Ts	**0** (0)	**1** (2.44)
200	A>G	A	Ts	**2** (4.08)	**6** (14.63)	**16227**	A>G	A	Ts	**2** (4.08)	**1** (2.44)
204	T>C	T	Ts	**4** (8.16)	**2** (4.88)	**16234**	C>T	C	Ts	**2** (4.08)	**2** (4.88)
227	A>G	A	Ts	**1** (2.04)	**0** (0)	**16235**	A>G	A	Ts	**0** (0)	**2** (4.88)
228	G>A	G	Ts	**1** (2.04)	**0** (0)	**16239**	C>T	C	Ts	**1** (2.04)	**0** (0)
234	A>G	A	Ts	**1** (2.04)	**2** (4.88)	**16245**	C>T	C	Ts	**1** (2.04)	**0** (0)
235	A>G	A	Ts	**3** (6.12)	**6** (14.63)	**16246**	A>T	A	Tv	**3** (6.12)	**0** (0)
247	G>A	G	Ts	**1** (2.04)	**0** (0)	**16248**	C>T	C	Ts	**1** (2.04)	**0** (0)
248	A DEL	A	Del	**15** (30.61)	**6** (14.63)	**16249**	T>C	T	Ts	**2** (4.08)	**0** (0)
249	A DEL	A	Del	**2** (4.08)	**3** (7.32)	**16258**	A DEL	A	Del	**4** (8.16)	**4** (9.76)
262	C>T	C	Ts	**1** (2.04)	**0** (0)	**16260**	C>T	C	Ts	**5** (10.20)	**2** (4.88)
263	A>G	A	Ts	**49** (100)	**41** (100)	**16261**	C>T	C	Ts	**0** (0)	**2** (4.88)
297	A>G	A	Ts	**1** (2.04)	**0** (0)	**16264**	C>T	C	Ts	**0** (0)	**1** (2.44)
302	A>ACC	A	Int	**1** (2.04)	**0** (0)	**16266**	C>T	C	Ts	**3** (6.12)	**0** (0)
310	T>C	T	Ts	**22** (44.90)	**23** (56.10)	**16270**	C>T	C	Ts	**1** (2.04)	**0** (0)
310	T>TCC	T	Int	**3** (6.12)	**2** (4.88)	**16271**	T>C	T	Ts	**1** (2.04)	**0** (0)
315	C>CC	C	Int	**25** (51.02)	**17** (41.46)	**16272**	A>G	A	Ts	**4** (8.16)	**4** (9.76)
329	G>A	G	Ts	**2** (4.08)	**0** (0)	**16274**	G>A	G	Ts	**1** (2.04)	**0** (0)
338	C>T	C	Ts	**1** (2.04)	**0** (0)	**16278**	C>T	C	Ts	**2** (4.08)	**2** (4.88)
373	A>G	A	Ts	**0** (0)	**1** (2.44)	**16287**	C>T	C	Ts	**1** (2.04)	**0** (0)
382	C>T	C	Ts	**2** (4.08)	**0** (0)	**16288**	T>C	T	Ts	**2** (4.08)	**0** (0)

Locus	Base Change	Ref Seq	Mut Type	Case n (%)	Control n (%)	Locus	Base Change	Ref Seq	Mut Type	Case n (%)	Control n (%)
16051	A>G	A	Ts	**3** (6.12)	**1** (2.44)	**16289**	A>G	A	Ts	**3** (6.12)	**1** (2.44)
16065	G>A	G	Ts	**1** (2.04)	**0** (0)	**16290**	C>T	C	Ts	**3** (6.12)	**7** (17.07)
16086	T>C	T	Ts	**2** (4.08)	**2** (4.88)	**16294**	C>T	C	Ts	**0** (0)	**1** (2.44)
16092	T>C	T	Ts	**0** (0)	**4** (9.76)	**16298**	T>C	T	Ts	**4** (8.16)	**3** (7.32)
16093	T>C	T	Ts	**3** (6.12)	**1** (2.44)	**16301**	C>T	C	Ts	**2** (4.08)	**0** (0)
16111	C>T	C	Ts	**1** (2.04)	**1** (2.44)	**16304**	T>C	T	Ts	**19** (38.78)	**8** (19.51)
16126	T>C	T	Ts	**0** (0)	**1** (2.44)	**16311**	T>C	T	Ts	**14** (28.57)	**5** (12.20)
16129	G>A	G	Ts	**5** (10.20)	**4** (9.76)	**16316**	A>G	A	Ts	**1** (2.04)	**0** (0)
16136	T>C	T	Ts	**3** (6.12)	**1** (2.44)	**16318**	A>G	A	Ts	**1** (2.04)	**0** (0)
16140	T>C	T	Ts	**1** (2.04)	**0** (0)	**16319**	G>A	G	Ts	**3** (6.12)	**7** (17.07)
16145	G>A	G	Ts	**5** (10.20)	**0** (0)	**16324**	T>C	T	Ts	**3** (6.12)	**0** (0)
16147	C>T	C	Ts	**0** (0)	**2** (4.88)	**16325**	T>C	T	Ts	**2** (4.08)	**0** (0)
16158	A>G	A	Ts	**1** (2.04)	**1** (2.44)	**16352**	T>C	T	Ts	**1** (2.04)	**0** (0)
16162	A>G	A	Ts	**3** (6.12)	**3** (7.32)	**16355**	C>T	C	Ts	**3** (6.12)	**0** (0)
16162	A DEL	A	Del	**0** (0)	**1** (2.44)	**16356**	T>C	T	Ts	**1** (2.04)	**1** (2.44)
16167	C>T	C	Ts	**1** (2.04)	**0** (0)	**16362**	T>C	T	Ts	**16** (32.65)	**15** (36.59)
16168	C>T	C	Ts	**0** (0)	**1** (2.44)	**16381**	T>C	T	Ts	**0** (0)	**1** (2.44)
16169	C>T	C	Ts	**1** (2.04)	**0** (0)	**16391**	G>A	G	Ts	**3** (6.12)	**0** (0)
16172	T>C	T	Ts	**4** (8.16)	**4** (9.76)	**16399**	A>G	A	Ts	**0** (0)	**2** (4.88)
16181	A>AC	A	Int	**1** (2.04)	**0** (0)	**16519**	T>C	T	Ts	**35** (71.43)	**24** (58.54)
16182	A>C	A	Tv	**1** (2.04)	**0** (0)	**16526**	G>A	G	Ts	**3** (6.12)	**2** (4.88)

CO1

There were more variations per sample among case groups than control groups. A total of 20 variations was found with 15 synonymous, 4 non-synonymous and 1 frameshift mutation (Table 37).

Haplogroups

There were 13 macrohaplogroups, the N Asian lineage A, B, F, H, N, P, R, U was 28 cases and 20 controls; the M Eurasian lineage D, E, G, M, Z was 21 in both groups (Figure 10 and 11). Haplogroups F (21.1%), M (17.8%), D (16.7%) were the most common among the groups. The frequency of haplogroup F (24.5%), M (20.4%) and R (18.4%) was higher in patients with breast cancer compared to the healthy volunteers.

Table 37. Comparison of variations found in CO1 region of case and control groups.

Mutation	Codon No	Codon Change	AA Change	PolyPhen2_ HumDiv_Pred	PP2_Hum Div_Pred	PolyPhen2_ HumVar_Pred	PP2_Hum Var_Pred
C5936T	11	AAC >AAT	Syn: Asn>Asn	Prob damaging	1.000	Prob damaging	1.000
C5939T	12	CAC>CAT	Syn: His>His	Prob damaging	0.997	Prob damaging	0.994
C5945T	14	GAC>GAT	Syn: Asp>Asp	Prob damaging	0.993	Prob damaging	0.99
T5964C	21	TTA>CTA	Syn: Leu>Leu	Benign	0.036	Benign	0.089
G5973A	24	GCA>ACA	Non-syn: Ala>Thr	Poss damaging	0.899	Poss damaging	0.502
T6011C	36	CTT>CTC	Syn: Leu>Leu	Poss damaging	0.953	Prob damaging	0.961
A6116G	71	ATA>ATG	Syn: Met>Met	-	-	-	-
A6120G	73	ATC>GTC	Non-syn: Ile>Val	Poss damaging	0.868	Prob damaging	0.925
C6209Del	102	TTC>TTd	Frameshift	-	-	-	-
T6216C	105	TTA>CTA	Syn: Leu>Leu	Prob damaging	1.000	Prob damaging	0.999
C6291T	130	CCT>TCT	Non-syn: Pro>Ser	Prob damaging	1.000	Prob damaging	0.997
T6293C	130	CCT>CCC	Syn: Pro>Pro	-	-	-	-
G6305T	134	GGG>GGT	Syn: Gly>Gly	-	-	-	-
T6392C	163	AAT>AAC	Syn: Asn>Asn	Benign	0.011	Benign	0.026
A6437G	178	CAA>CAG	Syn: Gln>Gln	Prob damaging	0.997	Prob damaging	0.995
G6480A	193	GTC>ATC	Non-syn: Val>Ile	Benign	0.000	Benign	0.001
T6524C	207	ACT>ACC	Syn: Thr>Thr	-	-	-	-
A6530G	209	CTA>CTG	Syn: Leu>Leu	-	-	-	-
A6599G	232	CAA>CAG	Syn: Gln>Gln	Prob damaging	1.000	Prob damaging	0.999
C6617T	238	TTC>TTT	Syn: Phe>Phe	Prob damaging	0.996	Prob damaging	0.990

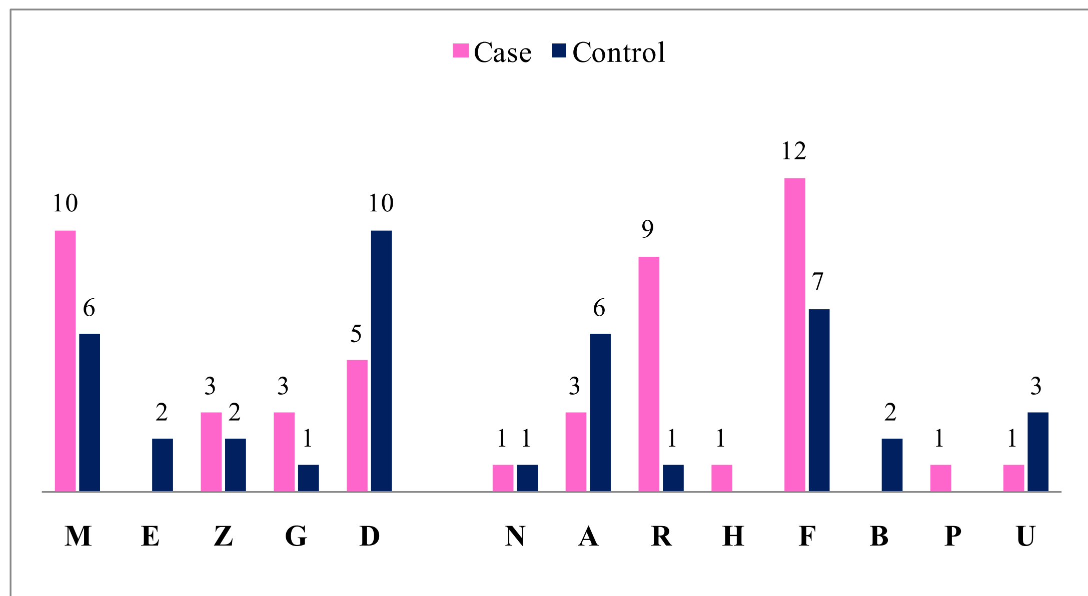


Figure 10. Comparison of haplogroups found in case and control groups.

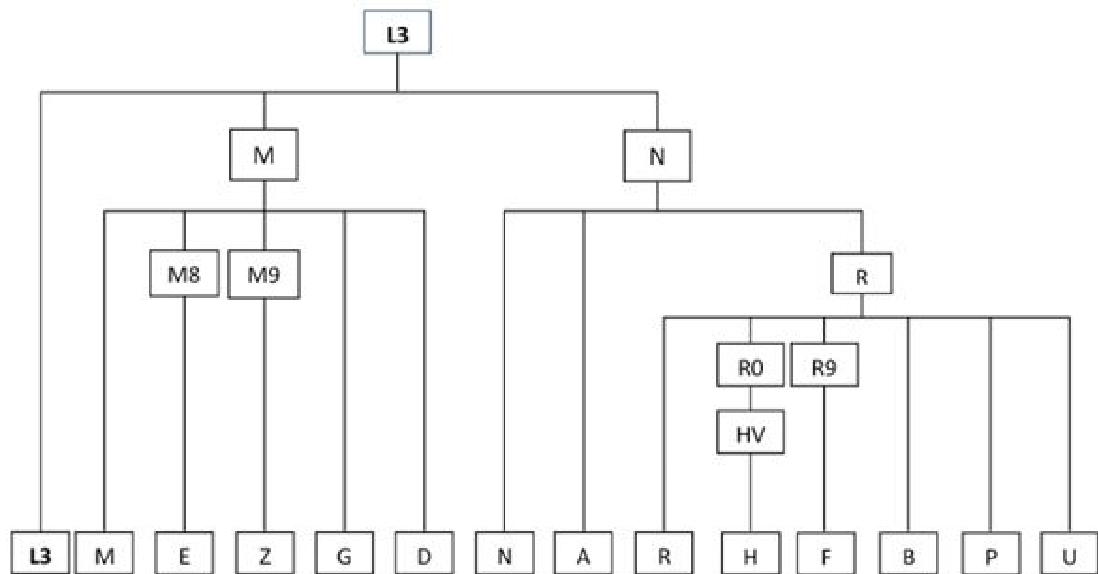


Figure 11. Branching of haplogroups from ancestral root (L3).

Table 38. Association of haplogroups with case and control groups.

Haplogroup	Case (%)	Control (%)	Total (%)	*P* value	OR (95% CI)
A	3 (6.1)	6 (14.6)	9 (10.0)	0.090	0.500 (0.225-1.113)
B	0 (0.0)	2 (4.9)	2 (2.2)	-	-
D	5 (10.2)	10 (24.4)	15 (16.7)	0.028	0.794 (0.646-0.976)
E	0 (0.0)	2 (4.9)	2 (2.2)	-	-
F	12 (24.5)	7 (17.1)	19 (21.1)	0.050	1.114 (1.000-1.240)
G	3 (6.1)	1 (2.4)	4 (4.4)	0.099	1.201 (0.966-1.493)
H	1 (2.0)	0 (0.0)	1 (1.1)	-	-
M	10 (20.4)	6 (14.6)	16 (17.8)	0.087	1.066 (0.991-1.147)
N	1 (2.0)	1 (2.4)	2 (2.2)	-	-
P	1 (2.0)	0 (0.0)	1 (1.1)	-	-
R	9 (18.4)	1 (2.4)	10 (11.1)	0.000	1.221 (1.096-1.361)
U	1 (2.0)	3 (7.3)	4 (4.4)	0.099	0.913 (0.818-1.017)
Z	3 (6.1)	2 (4.9)	5 (5.6)	0.442	1.032 (0.953-1.117)

Comparison of Major Tribes in the Study

Since the study was just for Mizo population, mother's and father's tribe was recorded. All the different sub-tribes and family names were grouped into the 7 major tribes (Chawngkunga, 1996; Thangluaia, 2012) as shown

in Figure 12. For d-loop and CO1 sequences analysis, 90 samples were analysed. The tree topology revealed a lack of monophyly within the case and the control samples, and also the different tribes are not monophyletic..

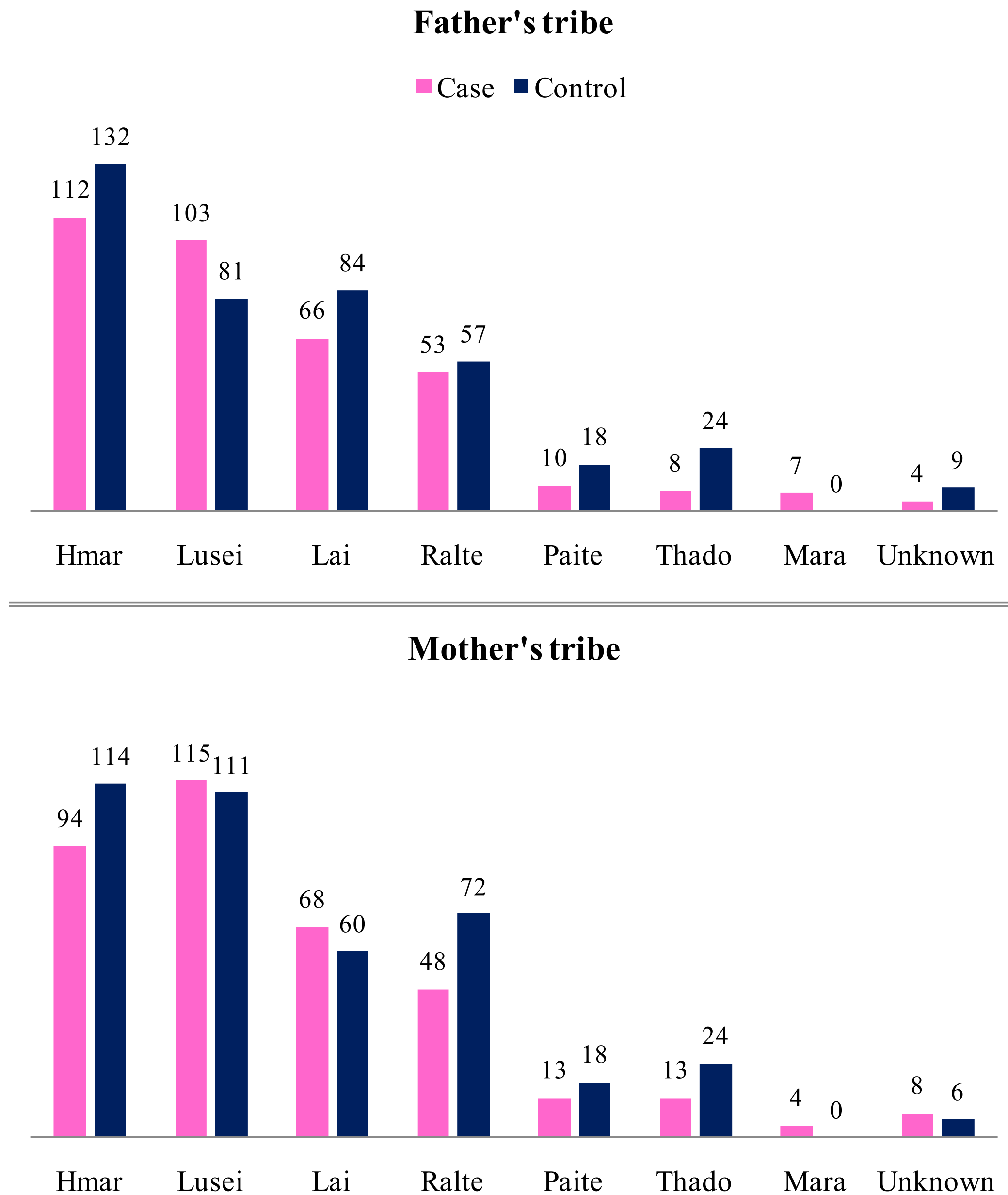


Figure 12. Comparison of major tribes in the study.

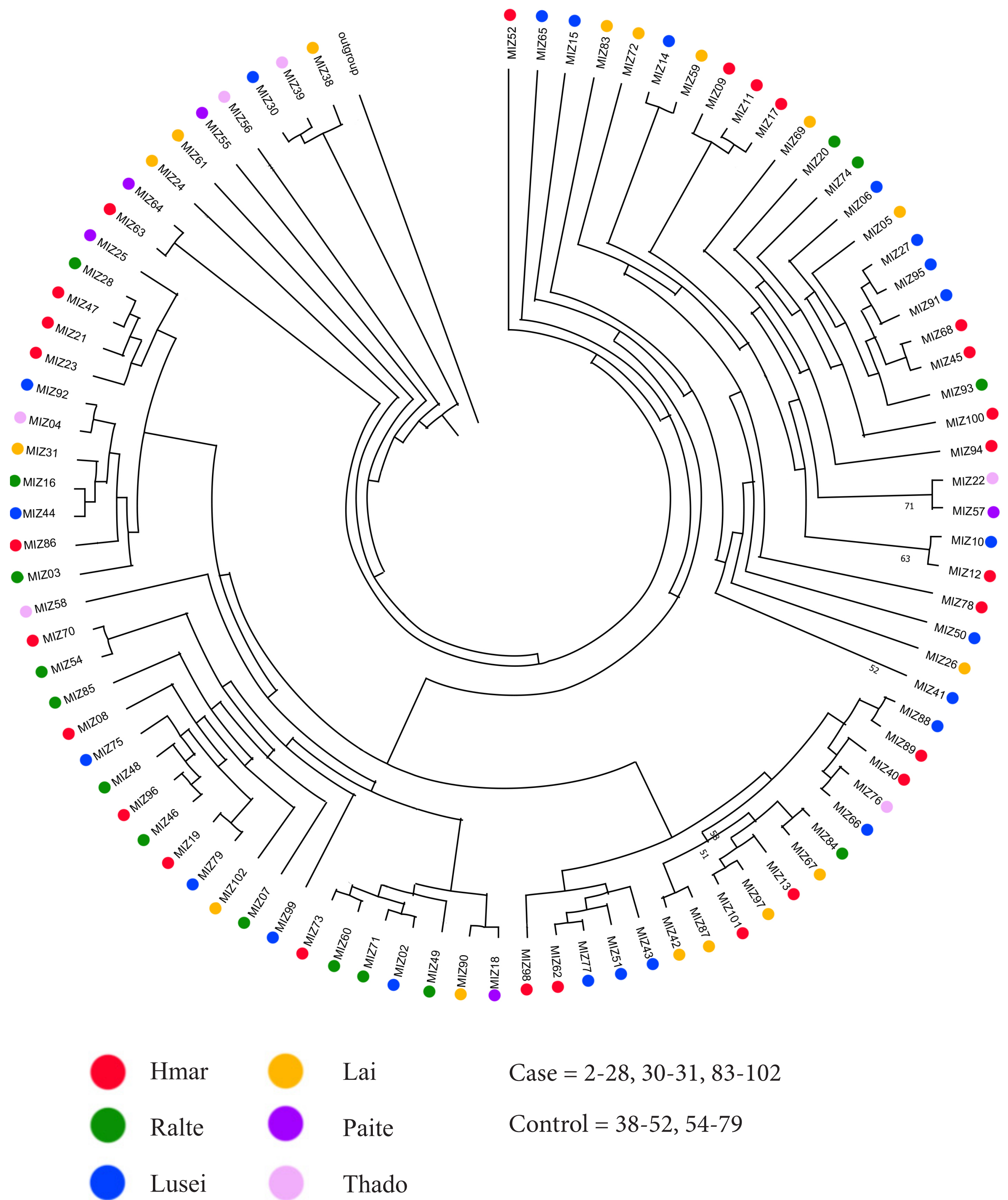


Figure 13. Phylogenetic tree based on CO1 sequences.

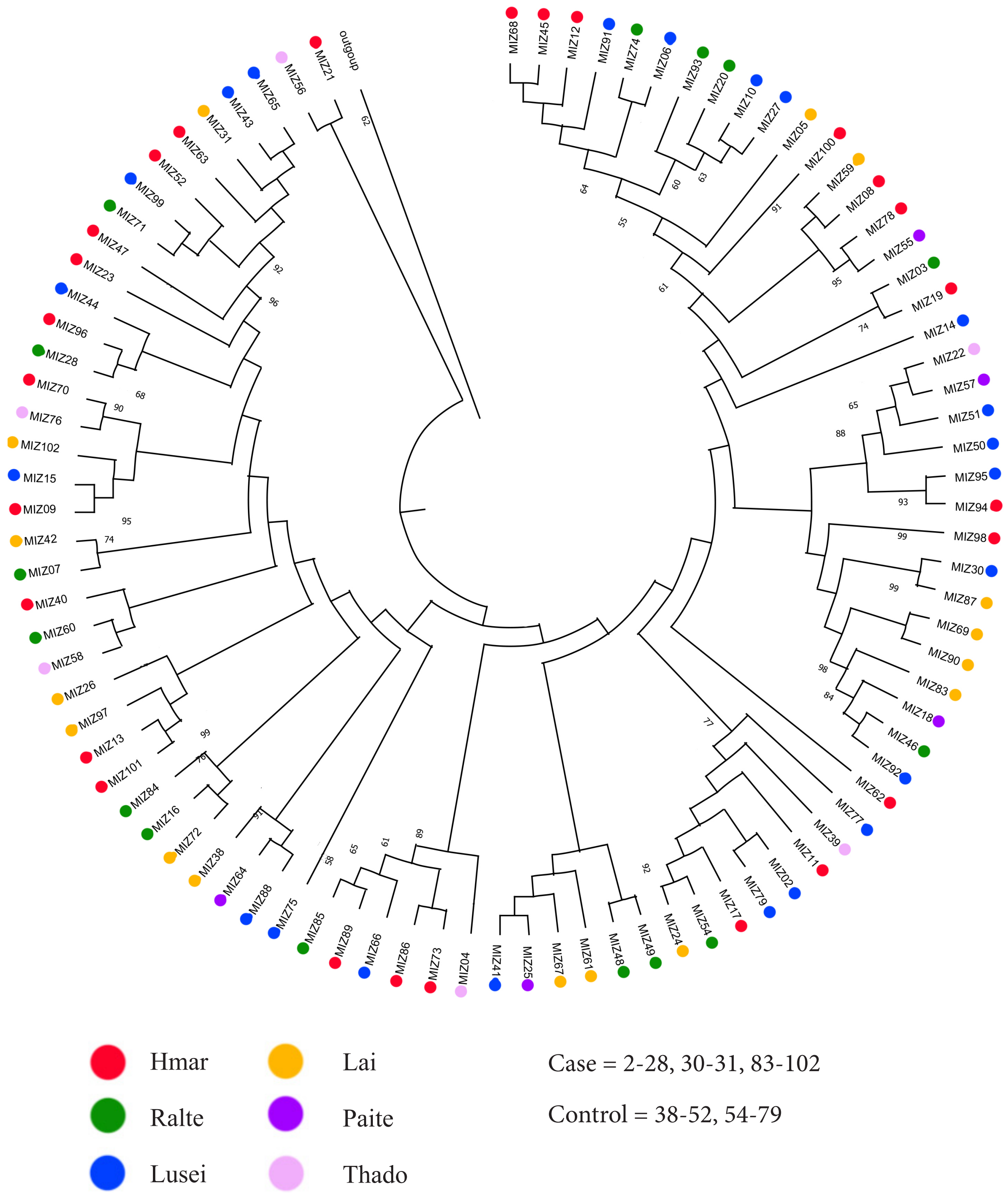


Figure 14. Phylogenetic tree based on d-loop sequences.

Association of mitochondrial DNA and breast cancer risk stratified by selected variables

All the variations were not analyzed for association with risk factors, only variations with significant difference between the groups were selected except for mutation T152C (Table 39). T152C was included because this mutation was found to have predisposition to breast cancer and was found in large numbers in both the groups. Of the 129 loci, only 9 loci were further selected, 6 variations from d-loop and 3 from CO1. The 9 loci were further analyzed stratified on the menopausal status of the subjects using wild type as referent (Table 40, 41 and 42). There was no association when stratified on menopausal status. However, interaction of the 9 variations with some of the selected characteristics show significant association (Table 43, 44 and 45). Association of mutation T146C with chewing of betel quid and tobacco and sleep duration of 7 hours or less was highly significant (Table 43). Association of mutation T152C with cigarette smoking and disturbed sleep pattern was highly significant (Table 43). Association of mutation A248Del with chewing of betel quid and tobacco and disturbed sleep pattern was highly significant (Table 43).

Association of mutation C5945T with chewing of betel quid and tobacco, using *tuibur*, cigarette smoking, disturbed sleep pattern and having history of miscarriage was highly significant (Table 44). Association of mutation T6392C with chewing of betel quid and tobacco, disturbed sleep pattern and having history of miscarriage was highly significant (Table 44). Association of mutation C6617T with chewing of betel quid, tobacco and *tuibur*, cigarette smoking, disturbed sleep pattern and having history of miscarriage was highly significant (Table 44).

Association of mutation T16304C with chewing of betel quid and tobacco, cigarette smoking including passive smoking, sleep duration of 7 hours or less, disturbed sleep pattern and history of miscarriage was highly significant (Table 45). Association of mutation T16311C with chewing of betel quid and tobacco, cigarette

smoking, sleep duration of 7 hours or less, disturbed sleep pattern and mother's age at birth of 35 years or over was highly significant (Table 45). Association of mutation T16519C with consumption of tobacco, cigarette smoking, disturbed sleep pattern and history of miscarriage was highly significant (Table 45).

Table 39. Comparison of mitochondrial variations with significance.

Mutation	Type	Hmtvar	Case (%)	Control (%)	χ2	*P* values
T146C	Regulatory	Predisposition	15 (30.61)	7 (17.07)	6.65	0.010
T152C	Regulatory	Predisposition	20 (40.82)	14 (34.15)	1.27	0.260
A248Del	Regulatory	-	15 (30.61)	6 (14.63)	9.56	0.002
C5945T	Coding	Prob damaging	44 (89.80)	24 (58.54)	35.43	0.000
T6392C	Coding	Deleterious variant	12 (24.49)	5 (12.20)	6.61	0.010
C6617T	Coding	Prob damaging	47 (95.92)	28 (68.29)	36.80	0.000
T16304C	Regulatory	Predisposition	19 (38.78)	8 (19.51)	11.83	0.001
T16311C	Regulatory	Predisposition	14 (28.57)	5 (12.20)	10.78	0.001
T16519C	Regulatory	Predisposition	35 (71.43)	24 (58.54)	4.93	0.026

Table 40. Association of mitochondrial variation at locus 146, 152 and 248 with breast cancer on menopausal status.

	T146C		**T152C**		**A248Del**	
	Wild	Mutant	Wild	Mutant	Wild	Mutant
Total						
Case	34 (69.4)	15 (30.6)	29 (59.2)	20 (40.8)	34 (69.4)	15 (30.6)
Control	34 (82.9)	7 (17.1)	27 (65.9)	14 (34.1)	35 (85.4)	6 (14.6)
OR (95% CI)	1.00	0.47 (0.26-0.84)	1.00	0.75 (0.46-1.24)	1.00	0.39 (0.21-0.72)
χ2 *P*	0.010		0.260		0.002	
Premenopausal						
Case	17 (65.4)	9 (34.6)	18 (69.2)	8 (30.8)	19 (73.1)	7 (26.9)
Control	22 (88.0)	3 (12.0)	16 (64.0)	9 (36.0)	23 (92.0)	2 (8.0)
OR (95% CI)	1.00	0.26 (0.11-0.60)	1.00	1.27 (0.65-2.48)	1.00	0.24 (0.09-0.62)
χ2 *P*	0.001		0.493		0.002	
Postmenopausal						
Case	17 (73.9)	6 (26.1)	11 (47.8)	12 (52.2)	15 (65.2)	8 (34.8)
Control	12 (75.0)	4 (25.0)	11 (68.8)	5 (31.3)	12 (75.0)	4 (25.0)
OR (95% CI)	1.00	0.94 (0.41-2.20)	1.00	0.42 (0.19-0.90)	1.00	0.63 (0.28-1.42)
χ2 *P*	0.895		0.025		0.259	

Table 41. Association of mitochondrial variation at locus 5945, 6392 and 6617 with breast cancer on menopausal status.

	C5945T		T6392C		C6617T	
	Wild	Mutant	Wild	Mutant	Wild	Mutant
Total						
Case	5 (10.2)	44 (89.8)	37 (75.5)	12 (24.5)	2 (4.1)	47 (95.9)
Control	17 (41.5)	24 (58.5)	36 (87.8)	5 (12.2)	13 (31.7)	28 (68.3)
OR (95% CI)	1.00	0.16 (0.08-0.31)	1.00	0.43 (0.22-0.83)	1.00	0.09 (0.04-0.23)
$\chi2$ *P*	0.000		0.010		0.000	
Premenopausal						
Case	4 (15.4)	22 (84.6)	19 (73.1)	7 (26.9)	1 (3.8)	25 (96.2)
Control	11 (44.0)	14 (56.0)	22 (88.0)	3 (12.0)	9 (36.0)	16 (64.0)
OR (95% CI)	1.00	0.23 (0.11-0.50)	1.00	0.37 (0.16-0.87)	1.00	0.07 (0.02-0.25)
$\chi2$ *P*	0.000		0.020		0.000	
Postmenopausal						
Case	1 (4.3)	22 (95.7)	18 (78.3)	5 (21.7)	1 (4.3)	22 (95.7)
Control	6 (37.5)	10 (62.5)	14 (87.5)	2 (12.5)	4 (25.0)	12 (75.0)
OR (95% CI)	1.00	0.08 (0.02-0.28)	1.00	0.51 (0.18-1.44)	1.00	0.14 (0.04-0.51)
$\chi2$ *P*	0.000		0.200		0.001	

Table 42. Association of mitochondrial variation at locus 16304, 16311 and 16519 with breast cancer on menopausal status.

	T16304C		T16311C		T16519C	
	Wild	Mutant	Wild	Mutant	Wild	Mutant
Total						
Case	30 (61.2)	19 (38.8)	35 (71.4)	14 (28.6)	14 (28.6)	35 (71.4)
Control	33 (80.5)	8 (19.5)	36 (87.8)	5 (12.2)	17 (41.5)	24 (58.5)
OR (95% CI)	1.00	0.38 (0.22-0.67)	1.00	0.35 (0.18-0.66)	1.00	0.56 (0.34-0.94)
$\chi2$ *P*	0.001		0.001		0.026	
Premenopausal						
Case	16 (61.5)	10 (38.5)	19 (73.1)	7 (26.9)	8 (30.8)	18 (69.2)
Control	21 (84.0)	4 (16.0)	23 (92.0)	2 (8.0)	12 (48.0)	13 (52.0)
OR (95% CI)	1.00	0.30 (0.14-0.66)	1.00	0.24 (0.09-0.62)	1.00	0.48 (0.25-0.93)
$\chi2$ *P*	0.002		0.002		0.029	
Postmenopausal						
Case	14 (60.9)	9 (39.1)	16 (69.6)	7 (30.4)	6 (26.1)	17 (73.9)
Control	12 (75.0)	4 (25.0)	13 (81.3)	3 (18.8)	5 (31.3)	11 (68.8)
OR (95% CI)	1.00	0.52 (0.23-1.17)	1.00	0.53 (0.22-1.28)	1.00	0.78 (0.34-1.75)
$\chi2$ *P*	0.111		0.155		0.542	

Table 43. Association of mitochondrial variations at locus 146, 152 and 248 with breast cancer risk.

		T146C		T152C		A248Del	
Variables	Type	Ca|Co	OR (95% CI)	Ca|Co	OR (95% CI)	Ca|Co	OR (95% CI)
Betel quid							
Never	Wild	09|06	1.00	05|04	1.00	06|04	1.00
Ever	Wild	25|28	0.94 (0.81-1.10)	24|23	1.02 (0.87-1.20)	28|31	0.95 (0.82-1.10)
Never	Mutant	02|00	1.00	06|02	3.00 (1.19-7.56)	05|02	2.50 (0.97-6.44)
Ever	Mutant	13|07	1.36 (1.05-1.78)	14|12	1.08 (0.86-1.35)	10|04	1.58 (1.13-2.21)
Tobacco							
Never	Wild	10|15	1.00	01|10	1.00	06|15	1.00
Ever	Wild	24|19	1.12 (0.94-1.34)	28|17	1.28 (1.08-1.53)	28|20	1.18 (1.00-1.40)
Never	Mutant	01|01	1.00	10|06	1.67 (0.93-2.99)	05|01	5.00 (1.45-17.27)
Ever	Mutant	14|06	1.53 (1.16-2.01)	10|08	1.12 (0.85-1.46)	10|05	1.41 (1.04-1.93)
Smoking							
Never	Wild	19|26	1.00	17|18	1.00	20|24	1.00
Ever	Wild	15|08	1.37 (1.07-1.75)	12|09	1.15 (0.90-1.48)	14|11	1.13 (0.90-1.42)
Never	Mutant	10|04	2.50 (1.28-4.88)	12|12	1.00	09|06	1.50 (0.83-2.72)
Ever	Mutant	05|03	1.29 (0.85-1.95)	08|02	2.00 (1.28-3.13)	06|00	1.00
PSmoking							
No	Wild	15|10	1.00	13|09	1.00	13|10	1.00
Yes	Wild	19|24	0.89 (0.75-1.06)	16|18	0.94 (0.78-1.15)	21|25	0.92 (0.78-1.08)
No	Mutant	06|02	3.00 (1.19-7.56)	08|03	2.67 (1.24-5.74)	08|02	4.00 (1.64-9.79)
Yes	Mutant	09|05	1.34 (0.98-1.84)	12|11	1.04 (0.82-1.32)	07|04	1.32 (0.93-1.89)
Tuibur							
Never	Wild	19|25	1.00	16|19	1.00	19|25	1.00
Ever	Wild	15|09	1.29 (1.02-1.64)	13|08	1.27 (0.99-1.64)	15|10	1.22 (0.97-1.54)
Never	Mutant	11|04	2.75 (1.42-5.32)	14|10	1.40 (0.88-2.24)	11|04	2.75 (1.42-5.32)
Ever	Mutant	04|03	1.15 (0.75-1.78)	06|04	1.22 (0.85-1.76)	04|02	1.41 (0.87-2.31)
Sleep duration							
≥8 hrs	Wild	19|31	1.00	11|05	1.00	14|04	1.00
≤7 hrs	Wild	15|03	0.78 (0.66-0.92)	18|22	0.90 (0.76-1.08)	20|31	0.80 (0.68-0.94)
≥8 hrs	Mutant	05|02	2.50 (0.97-6.44)	09|00	1.00	06|01	6.00 (1.77-20.37)
≤7 hrs	Mutant	10|05	1.41 (1.04-1.93)	11|14	0.89 (0.71-1.11)	09|05	1.34 (0.98-1.84)
Sleep pattern							
Undisturbed	Wild	14|24	1.00	16|17	1.00	17|24	1.00
Disturbed	Wild	20|10	1.41 (1.14-1.76)	13|10	1.14 (0.90-1.45)	17|11	1.24 (1.00-1.55)
Undisturbed	Mutant	10|05	2.00 (1.08-3.72)	08|12	0.67 (0.40-1.12)	07|05	1.40 (0.72-2.72)
Disturbed	Mutant	05|02	1.58 (0.98-2.54)	12|02	2.45 (1.59-3.77)	08|01	2.83 (1.55-5.15)
Miscarriage							
Never	Wild	20|23	1.00	17|19	1.00	20|25	1.00
Ever	Wild	14|11	1.13 (0.90-1.42)	12|08	1.22 (0.95-1.59)	14|10	1.18 (0.94-1.50)
Never	Mutant	10|04	2.50 (1.28-4.88)	13|08	1.62 (0.98-2.70)	10|02	5.00 (2.08-12.01)
Ever	Mutant	05|03	1.29 (0.85-1.95)	07|06	1.08 (0.79-1.48)	05|04	1.12 (0.76-1.63)
Maternal age							
≤34	Wild	27|25	1.00	21|20	1.00	26|28	1.00
≥35	Wild	07|09	0.88 (0.66-1.17)	08|07	1.07 (0.80-1.43)	08|07	1.07 (0.80-1.43)
≤34	Mutant	12|07	1.71 (1.00-2.94)	18|12	1.50 (0.98-2.29)	13|04	3.25 (1.70-6.21)
≥35	Mutant	03|00	1.00	02|02	1.00	02|02	1.00

Table 44. Association of mitochondrial variations at locus 5945, 6392 and 6617 with breast cancer risk.

		C5945T		T6392C		C6617T	
Variables	Type	Ca|Co	OR (95% CI)	Ca|Co	OR (95% CI)	Ca|Co	OR (95% CI)
Betel quid							
Never	Wild	01|04	1.00	08|05	1.00	01|01	1.00
Ever	Wild	04|13	0.55 (0.40-0.77)	29|31	0.97 (0.84-1.12)	01|12	0.29 (0.16-0.52)
Never	Mutant	10|02	5.00 (2.08-12.01)	03|01	3.00 (0.81-11.08)	10|05	2.00 (1.08-3.72)
Ever	Mutant	34|22	1.24 (1.06-1.45)	09|04	1.50 (1.07-2.11)	37|23	1.27 (1.09-1.47)
Tobacco							
Never	Wild	01|06	1.00	09|15	1.00	00|04	1.00
Ever	Wild	04|11	0.60 (0.43-0.84)	28|21	1.15 (0.98-1.36)	02|09	0.47 (0.30-0.73)
Never	Mutant	10|10	1.00	02|01	2.00 (0.50-8.00)	11|12	0.92 (0.57-1.47)
Ever	Mutant	34|14	1.56 (1.30-1.87)	10|04	1.58 (1.13-2.21)	36|16	1.50 (1.27-1.78)
Smoking							
Never	Wild	04|10	1.00	20|25	1.00	02|07	1.00
Ever	Wild	01|07	0.38 (0.21-0.69)	17|11	1.24 (1.00-1.55)	00|06	1.00
Never	Mutant	25|20	1.25 (0.89-1.76)	09|05	1.80 (0.96-3.38)	27|23	1.17 (0.85-1.62)
Ever	Mutant	19|04	2.18 (1.60-2.98)	03|00	1.00	20|05	2.00 (1.51-2.65)
PSmoking							
No	Wild	01|07	1.00	14|10	1.00	00|04	1.00
Yes	Wild	04|10	0.63 (0.45-0.88)	23|26	0.94 (0.80-1.11)	02|09	0.47 (0.30-0.73)
No	Mutant	20|05	4.00 (2.27-7.04)	07|02	3.50 (1.41-8.67)	21|08	2.62 (1.64-4.20)
Yes	Mutant	24|19	1.12 (0.94-1.34)	05|03	1.29 (0.85-1.95)	26|20	1.14 (0.96-1.35)
Tuibur							
Never	Wild	04|10	1.00	21|25	1.00	01|09	1.00
Ever	Wild	01|07	0.38 (0.21-0.69)	16|11	1.21 (0.97-1.51)	01|04	0.50 (0.27-0.94)
Never	Mutant	26|19	1.37 (0.97-1.93)	09|04	2.25 (1.14-4.44)	29|20	1.45 (1.04-2.01)
Ever	Mutant	18|05	1.90 (1.43-2.53)	03|01	1.73 (0.90-3.33)	18|08	1.50 (1.18-1.91)
Sleep duration							
≥8 hrs	Wild	02|02	1.00	16|04	1.00	00|02	1.00
≤7 hrs	Wild	03|15	0.45 (0.31-0.64)	21|32	0.81 (0.69-0.95)	02|11	0.43 (0.28-0.66)
≥8 hrs	Mutant	18|03	6.00 (2.96-12.15)	04|01	4.00 (1.13-14.17)	20|03	6.67 (3.31-13.43)
≤7 hrs	Mutant	26|21	1.11 (0.94-1.31)	08|04	1.41 (1.00-2.00)	27|25	1.04 (0.89-1.22)
Sleep pattern							
Undisturbed	Wild	03|11	1.00	18|25	1.00	01|07	1.00
Disturbed	Wild	02|06	0.58 (0.36-0.92)	19|11	1.31 (1.06-1.63)	01|06	0.41 (0.22-0.75)
Undisturbed	Mutant	21|18	1.17 (0.81-1.68)	06|04	1.50 (0.72-3.11)	23|22	1.05 (0.75-1.47)
Disturbed	Mutant	23|06	1.96 (1.51-2.54)	06|01	2.45 (1.33-4.51)	24|06	2.00 (1.54-2.59)
Miscarriage							
Never	Wild	04|10	1.00	22|24	1.00	02|09	1.00
Ever	Wild	01|07	0.38 (0.21-0.69)	15|12	1.12 (0.90-1.39)	00|04	1.00
Never	Mutant	26|17	1.53 (1.07-2.18)	08|03	2.67 (1.24-5.74)	28|18	1.56 (1.11-2.19)
Ever	Mutant	18|07	1.60 (1.25-2.06)	04|02	1.41 (0.87-2.31)	19|10	1.38 (1.11-1.72)
Maternal age							
≤34	Wild	05|15	1.00	29|28	1.00	02|11	1.00
≥35	Wild	00|02	1.00	08|08	1.00	00|02	1.00
≤34	Mutant	34|17	2.00 (1.43-2.80)	10|04	2.50 (1.28-4.88)	37|21	1.76 (1.29-2.40)
≥35	Mutant	10|07	1.20 (0.90-1.58)	02|01	1.41 (0.71-2.83)	10|07	1.20 (0.90-1.58)

Table 45. Association of mitochondrial variations at locus 16304, 16311 and 16519 with breast cancer risk.

		T16304C		T16311C		T16519C	
Variables	Type	Ca|Co	OR (95% CI)	Ca|Co	OR (95% CI)	Ca|Co	OR (95% CI)
Betel quid							
Never	Wild	06|04	1.00	08|05	1.00	00|01	1.00
Ever	Wild	24|29	0.91 (0.78-1.06)	27|31	0.93 (0.80-1.08)	14|16	0.94 (0.76-1.15)
Never	Mutant	05|02	2.50 (0.97-6.44)	03|01	3.00 (0.81-11.08)	11|05	2.20 (1.19-4.05)
Ever	Mutant	14|06	1.53 (1.16-2.01)	11|04	1.66 (1.19-2.31)	24|19	1.12 (0.94-1.34)
Tobacco							
Never	Wild	06|14	1.00	08|15	1.00	04|08	1.00
Ever	Wild	24|19	1.12 (0.94-1.34)	27|21	1.13 (0.96-1.34)	10|09	1.05 (0.81-1.37)
Never	Mutant	05|02	2.50 (0.97-6.44)	03|01	3.00 (0.81-11.08)	07|08	0.88 (0.49-1.57)
Ever	Mutant	14|06	1.53 (1.16-2.01)	11|04	1.66 (1.19-2.31)	28|16	1.32 (1.11-1.58)
Smoking							
Never	Wild	16|23	1.00	22|26	1.00	09|14	1.00
Ever	Wild	14|10	1.18 (0.94-1.50)	13|10	1.14 (0.90-1.45)	05|03	1.29 (0.85-1.95)
Never	Mutant	13|07	1.86 (1.09-3.16)	07|04	1.75 (0.86-3.56)	20|16	1.25 (0.86-1.83)
Ever	Mutant	06|01	2.45 (1.33-4.51)	07|01	2.65 (1.44-4.84)	15|08	1.37 (1.07-1.75)
PSmoking							
No	Wild	13|08	1.00	15|11	1.00	06|05	1.00
Yes	Wild	17|25	0.82 (0.69-0.99)	20|25	0.89 (0.75-1.06)	08|12	0.82 (0.63-1.06)
No	Mutant	08|04	2.00 (1.00-4.00)	06|01	6.00 (1.77-20.37)	15|07	2.14 (1.28-3.60)
Yes	Mutant	11|04	1.66 (1.19-2.31)	08|04	1.41 (1.00-2.00)	20|17	1.08 (0.90-1.31)
Tuibur							
Never	Wild	15|23	1.00	19|27	1.00	07|14	1.00
Ever	Wild	15|10	1.22 (0.97-1.54)	16|09	1.33 (1.05-1.69)	07|03	1.53 (1.03-2.26)
Never	Mutant	15|06	2.50 (1.45-4.32)	11|02	5.50 (2.30-13.13)	23|15	1.53 (1.05-2.23)
Ever	Mutant	04|02	1.41 (0.87-2.31)	03|03	1.00	12|09	1.15 (0.90-1.48)
Sleep duration							
≥8 hrs	Wild	13|04	1.00	15|03	1.00	07|00	1.00
≤7 hrs	Wild	17|29	0.77 (0.64-0.91)	20|33	0.78 (0.66-0.91)	07|17	0.64 (0.50-0.83)
≥8 hrs	Mutant	07|01	7.00 (2.09-23.47)	05|02	2.50 (0.97-6.44)	13|05	2.60 (1.43-4.72)
≤7 hrs	Mutant	12|07	1.31 (1.00-1.71)	09|03	1.73 (1.19-2.53)	22|19	1.08 (0.90-1.28)
Sleep pattern							
Undisturbed	Wild	14|23	1.00	17|26	1.00	05|12	1.00
Disturbed	Wild	16|10	1.26 (1.01-1.59)	18|10	1.34 (1.07-1.68)	09|05	1.34 (0.98-1.84)
Undisturbed	Mutant	10|06	1.67 (0.93-2.99)	07|03	2.33 (1.07-5.09)	19|17	1.12 (0.77-1.63)
Disturbed	Mutant	09|02	2.12 (1.36-3.30)	07|02	1.87 (1.19-2.94)	16|07	1.51 (1.17-1.95)
Miscarriage							
Never	Wild	19|23	1.00	19|24	1.00	10|12	1.00
Ever	Wild	11|10	1.05 (0.82-1.34)	16|12	1.15 (0.93-1.43)	04|05	0.89 (0.61-1.31)
Never	Mutant	11|04	2.75 (1.42-5.32)	11|03	3.67 (1.75-7.66)	20|15	1.33 (0.91-1.96)
Ever	Mutant	08|04	1.41 (1.00-2.00)	03|02	1.22 (0.73-2.05)	15|09	1.29 (1.02-1.64)
Maternal age							
≤34	Wild	23|26	1.00	30|28	1.00	12|16	1.00
≥35	Wild	07|07	1.00	05|08	0.79 (0.57-1.09)	02|01	1.41 (0.71-2.83)
≤34	Mutant	16|06	2.67 (1.55-4.58)	09|04	2.25 (1.14-4.44)	27|16	1.69 (1.18-2.41)
≥35	Mutant	03|02	1.22 (0.73-2.05)	05|01	2.24 (1.20-4.16)	08|08	1.00

CHAPTER 6

DISCUSSION

6. Discussion

Countless number of studies have been conducted in different parts of the world using case control method since the time Janet Elizabeth Lane-Claypon carried out the first modern case control study in United Kingdom in 1926. Studies have drawn contrasting conclusions to the etiology of breast cancer as to pinpoint specific factors. This could be due to differences in sample sizes and races of study populations, local customs, lifestyle and culture, and health care conditions. In spite of contrasting conclusions, there are solid consensus on some of these factors. According to American Cancer Society, the most important non-modifiable risk factors are age, gender, genetic predisposition, a family history of breast cancer, personal history of breast cancer, benign breast disease, and lifetime menstrual cycles. Modifiable risk factors are consumption of alcohol, being overweight or obese after menopause, not being physical active, nulliparity or late age at first birth, use of oral contraceptives and hormone therapy, and not breastfeeding. There are important controversial factors like miscarriage, use of antiperspirants and bras. There are also unclear factors like the effects of diet and vitamins, chemicals, tobacco smoke and night shift work on breast cancer risk.

The frustration of breast cancer epidemiology has been that the strongest risk factors – genetic susceptibility are rare and some of the most common risk factors – age, family history are not amenable to change. The entire proportion of explained population attributable risk from known factors ranges from 15-55 percent (Seidman *et al*., 1982). Many risk factors are inextricably tied to our modern lifestyle, and clearly there are etiological factors that remain unknown or controversial. Due to different conclusions made by different studies and differences seen geographically, a retrospective case control study too was carried out among Mizo female population to study which factors could influence risk. With this in mind, established as well as suspected risk factors were used to compare between breast cancer cases and normal healthy controls free of cancer with a ratio of 1:1.1.

6.1. Demographic Characteristics

Studies have shown a large difference in breast cancer rates between rural and urban regions, the reasons for these differences are not well understood. About 55 percent of the global burden is among developed countries and the main reason put forward was lifestyle factors. Studies in India found living in rural areas reduces risk (Nagrani *et al*., 2014) and that central adiposity and age at first childbirth to be the factors influencing risk (Nagrani *et al*., 2016).

However, in this study, there were more controls (74.8%) from urban areas and as such no conclusion could be made on the impact of area of occurrence. But residing in urban areas did influence BMI as more controls were overweight, more controls unmarried and later age at marriage. Among postmenopausal groups, risk was observed with being overweight or obese (OR = 2.145, 95% CI: 1.21-3.82). Association between Body mass index (BMI) and breast cancer risk differs by menopausal status in that high BMI may be associated with a lower risk of premenopausal breast cancer, but is strongly associated with a higher risk of postmenopausal breast cancer (World Cancer Research Fund and American Institute for Cancer Research, 2007). This could be because after menopause, ovarian estrogen production ceases and adipose tissue becomes a key endogenous source of circulating estrogens (Lorincz and Sukumar, 2006; Kendall *et al*., 2007).

6.2. Dietary Habits

When evaluating the relation between dietary habits and breast cancer, retrospective case control studies may afford important insights, but are susceptible to both recall bias and selection bias especially against fondness of the food. Affected individuals may associate their malignancy with foods perceived to be poor in nutritional value and over report them relative to unaffected controls. Prospective studies assess diet before breast cancer diagnosis and address some of the inherent limitations of

case control studies. But provide only a snapshot of dietary intake and is challenging to accurately assess an individual's diet over prolonged periods of time. Based on an extensive review of the literature, the World Cancer Research Fund classified as 'Limited evidence, no conclusion' the association of breast cancer with dietary fiber, vegetables and fruits, soya and soya products, meat, fish, milk and dairy products, folate, vitamin D, calcium, selenium glycemic index, and dietary patterns. The expert panel found 'Limited suggestive' association for total fat and postmenopausal but not premenopausal breast cancer (World Cancer Research Fund and American Institute for Cancer Research, 2010).

In this study, significant differences were observed between the groups in their dietary habits but only frequent consumption of fermented pork had any significant association with risk (OR = 2.228, 95% CI: 1.51-3.28) and among postmenopausal groups frequent consumption of sodium bicarbonate (OR = 2.457, 95% CI: 1.38-4.37). Interaction of GSTT1 and GSTM1 with dietary habits did not yield and significant association. Fermented pork or pork fat (*Sa-um*) is a unique delicacy of the Mizos. It is used in combination or separately with sodium bicarbonate in the preparation of foods (*bai* and *bawl*). The potential role of fat intake as a risk factor for breast cancer received widespread attention after a report of dramatic differences in fat consumption and breast cancer incidence across countries (Armstrong and Doll, 1975). High fat diets have been shown to induce mammary tumors in rodents (Fay and Freedman, 1997). Lowering dietary fat intake has also been associated with a decrease in estradiol levels (Wu *et al*., 1999). A pooled analysis of eight prospective cohorts found that total fat intake was not related to risk of either pre or postmenopausal breast cancer, but observed a modest association with saturated fat intake (Smith-Warner *et al*., 2001a). Different types of fat may have different effects on the breast. Olive oil, which has a high content of monosaturated fat, has been shown to lower breast cancer risk in some studies (Wolk *et al*., 1998; Voorrips *et al*., 2002). Fat intake in adulthood is

not consistently associated with breast cancer and fat intake restriction should not be considered as a primary preventive measure for breast cancer.

Some studies in breast cancer etiology have revealed additional complexities of the potential relevance of diet (Holmes *et al*., 2004; Kotepui, 2016). Despite a wide international variation in breast cancer rates, given that diet is a promising target for prevention, additional studies of diet with improved means of assessing exposures and relevant biomarkers are certainly warranted. A careful consideration of the study design and implementation and accuracy of dietary assessment is also required. Additionally, several determinants of breast cancer differ in premenopausal and postmenopausal women. Exposures occurring at various stages of life, from as early as in utero up to age at diagnosis can potentially have an important impact on breast cancer risk as the mammary tissue may be particularly susceptible to environmental influence at that time (Land *et al*., 2003).

6.3. Behavioral Habits

Smoking and physical inactivity are few behavioral risk factors amenable to change, and as such represents an opportunity to reduce the burden of disease from breast cancer. There is growing evidence that smoking may slightly increase the risk of breast cancer. In a meta-analysis, current and former smoking were weakly associated with breast cancer risk; a stronger association was observed in women who initiated smoking before first birth (Gaudet *et al*., 2013). Statistically significant effects have been seen for early age at starting, and for heavy, current, and passive smoking. In some studies, there is significant increase in breast cancer risk in never smokers with longterm exposure to passive smoking, while other scientist rejects the evidence of this association as inconsistent (Ambrosone and Shields, 1999; Wartenberg *et al*., 2000).

In this study, association with risk was seen with cigarette smoking (OR = 1.528, 95% CI: 1.15-2.03) especially in the postmenopausal groups (OR = 2.968, 95%

CI: 1.61-5.46) and had influence even on better survival. It was observed that non-smokers had a far better outcome than smokers (OR = 1.272, 95% CI: 1.085-1.491). The results would have been interesting if the age when they first started smoking was also recorded, but this too could be prone to difficulties in recollecting specific time when they first started smoking.

An analysis of interactions between alcohol and other risk factors indicates few significant interactions even among risk factors known to exhibit independent effects. But studies with independent effects such as parity, excess BMI, use of hormonal contraceptives or hormone replacement therapy did not exhibit interactions with alcohol (Collaborative Group on Hormonal Factors in Breast Cancer, 2002). While some studies have documented alcohol-HRT interactions for breast cancer; some are based on the absence of an HRT effect among women who do not drink or an alcohol effect only among never users of HRT (Colditz *et al*., 1990; Gapstur *et al*., 1992; Terry *et al*., 2006). In this study, interaction of alcohol with other independent risk factors did not exhibit any significant association.

The average annual level of physical activity over a woman's lifetime appears to be an important determinant of breast cancer risk. Although many details remain to be explained, there is sufficient evidence to conclude that a physically active lifestyle that persists over a woman's lifetime lowers her breast cancer risk, relative to an inactive woman. It is likely that physical activity is associated with decreased breast cancer risk via multiple interrelated biologic pathways that may involve adiposity, sex hormones, insulin resistance, adipokines, and chronic inflammation (Friedenreich *et al*., 2002). Inherent in the studies of physical activity is the difficulty in assessing usual physical activity over lifetime, adjustment for confounders and assessment of effect modification by other factors or characteristics. Some studies have utilized comprehensive assessments of lifetime physical activity, whereas others have used single-item measures. In this study, one difficulty was subjects were prone to recall

error, social desirability and other biases as well. But an overall picture was observed by calculating the metabolic equivalent task (MET) in a typical 24 hours duration.

Although observational studies cannot prove a causal relationship, when studies in different populations have similar results and when a possible mechanism for a causal relationship exists, this provides evidence of a causal connection. In this study, significant association was seen with heavy physical activity among premenopausal groups (OR = 4.515, 95% CI: 2.68-7.62) and that cases with moderate activity had better survival outcome (OR = 1.499, 95% CI: 1.051-2.139). Exercise has many biological effects on the body, some of which is lowering the levels of sex hormones, such as estrogen, and growth factors that have been associated with cancer development and progression. Also preventing high blood levels of insulin thereby increasing circulating sex hormone binding globulin which binds reversibly to estrogens to affect their bioavailability (Winzer *et al*., 2011). In addition, other research findings have raised the possibility that physical activity may have beneficial effects on survival for patients with breast, colorectal, and prostate cancers (Campbell *et al*., 2019; Schmitz *et al*., 2019).

6.4. Environmental Exposure and Medical History

There is little question that exposure to radiation is associated with an increased breast cancer risk. The challenge is to better understand which factors that characterize the nature of the radiation exposure are most important in determining the subsequent risk of breast cancer. Many methodological issues complicate the study of past exposure and methodologies for reconstructing doses of exposure are varied, complex, and imperfect. Even when dose is well measured, dose-response relationships are likely to be influenced by a variety of individual-level factors such as the timing of exposure, duration of exposure, and individual variation in susceptibility to the effects of exposure. While there is biological plausibility for associations between exposures and breast cancer in humans, epidemiologic evidence is currently limited. Evidence

for association is particularly strong for exposure to ionizing radiation but weak for exposure to environmental pollutants, extremely low frequency (ELF), trace elements, and heavy metals. In this study, there was no significant association to risk with the environmental pollutant that could be computed or observed by the subjects. If studies paying particular attention to dose rates, the timing of exposure, gene-environment interactions, and subgroup-specific effects could be performed instead of observational retrospective study, the role of all these environmental factors in relation to breast cancer risk would be better understood.

Hypertension increases the risk of all malignancies (Grossman *et al*., 2001) via hypothesized pathways relating to abnormalities of vascular smooth muscle proliferation, carcinogen binding to DNA, or angiogenesis (Felmeden and Lip, 2001). High levels of the protein GRK4 (G-protein coupled receptor kinase 4) that have been shown to cause hypertension may be implicated in breast cancer carcinogenesis (Sun *et al*., 2018). According to the American Heart Association, GRK4 expression is higher in breast cancer than in normal breast tissues and that two GRK4 variants have been identified as both hypertension and breast cancer risk loci (Yue *et al*., 2019). A meta-analysis of 30 studies, showed statistically significant association between hypertension and increased breast cancer risk and a positive association among postmenopausal women (Han *et al*., 2017). In recent study, a hospital-based case control among the Bangladeshi women showed significant association between breast cancer and hypertension (Islam *et al*., 2022). Treatment for hypertension has also been associated with breast cancer risk in several studies although the evidence is inconsistent (Grossman *et al*., 2001). In this study, there was strong association of breast cancer risk with hypertension (OR = 16.392, 95% CI: 3.30-81.54). But association could not be made when stratified on their menopausal status.

6.5. Menstrual History

Observation of age-incidence curves for breast cancer showed rates increase rapidly during the reproductive years and level off after menopause, which implicate the role of ovarian hormones estradiol and progesterone (MacMahon *et al*., 1973). Determining menopausal status can be complicated, the transition from premenopausal to postmenopausal stage is often several years in length, varies in symptomology and duration, and may not be measurable by menstrual patterns in women with a history of history of menopausal surgeries (i.e., hysterectomy, oophorectomy), or menopausal hormone therapy (HT) use. Given these complexities, there is no standardized definition for menopausal status in epidemiologic studies. In this study, a woman was considered postmenopausal – if she had undergone bilateral oophorectomy; if she affirmed that her menstruations had ceased for at least 6 consecutive months before diagnosis of breast cancer or prior to interview for control; if the above information were unavailable or inconclusive, 55 years and above. There was no significant association of age at menarche, but there was significant association with age at menopause (OR = 9.809, 95% CI: 4.42-21.79) and influence in survival was observed with lifetime cumulative number of menstrual cycles (OR = 1.761, 95% CI: 1.246-2.488). This could probably be due to prolonged exposure of breast epithelium to estrogen (Henderson *et al*., 1985)

6.6. Reproductive History

Reproductive history is the most studied factors and with the most contrasting results. A case control study in Morocco showed nulliparity was significantly related to an increased risk of breast cancer, and early age at first full-term pregnancy was associated with a decreased risk of breast cancer (Khalis *et al*., 2018). Another case control study in India found that early age at marriage, early age at first pregnancy, extended period of breastfeeding, and increased number of live births are protective for breast cancer (Babita *et al*., 2014). According to the American Cancer society, late age at first birth and nulliparity are modifiable risk factors, and that breastfeeding

may slightly lower risk of breast cancer. They have classified miscarriage under controversial factors. In this study, there was positive association with early age at first childbirth (OR = 2.73, 95% CI: 1.49-5.02) but no association was observed for number of live births or breastfeeding duration. The exact mechanism by which an early first birth protects against breast cancer remains incompletely understood, but has primarily been attributed to shortening of the time window of high susceptibility to carcinogenic transformations (Russo *et al*., 2000). Breast cancer risk appears to be related to timing of first birth and age at subsequent births. Higher parity and early age at first birth have both been associated with decreased lifetime incidence of breast cancer (Kelsey *et al*., 1993; Chie *et al*., 2000; Pathak *et al*., 2000). Studies have shown that risk is reduced to half in women having their first child before the age of 20 compared to those having their first child after the age of 30 (Russo *et al*., 2000). There is also some evidence that the interval between age at menarche and age at first birth may be relevant to breast cancer risk (Andrieu *et al*., 1998, 2000; Clavel-Chapelon and E3N Group, 2002; Li *et al*., 2008).

The use of oral contraceptive was not of much significance in this study as only the most recent generations of women have had the opportunity to accumulate long term pill use from younger age and most of the older generations had not practiced its use. In fact, oral contraceptive pills were first released in 1960 (Edgren, 1991). But in India, it was launched in 1987 with the brand name Mala D (National Health Portal, 2015). Results from other studies on the role of oral contraceptives to breast cancer proneness have been somewhat conflicting too. But, data from 54 studies concluded that current use of oral contraceptives poses a slight (24%) increase in the risk, which disappears 10 years after the cessation of use (Collaborative Group on Hormonal Factors in Breast Cancer, 1996).

6.7. Association of Risk Factors with Glutathione S-transferase

Large inter-ethnic differences have been reported in the frequencies of the GSTM1 and GSTT1 null genotype. The frequency of the GSTM1 null genotype is around 50 percent in Caucasians and Asians, and only 27 percent in Africans (Garte *et al*., 2001). Whereas, the frequency of GSTT1 null genotype is around 60 percent in Asians and 20 percent in Caucasians (Nelson *et al*., 1995). Studies thus show contrasting results but these deletion variants have been useful for molecular epidemiological studies of cancer because they divide study subjects into two well-defined susceptibility classes: those who are and those who are not able to detoxify potential carcinogens by the metabolic pathways regulated by GSTM1 and GSTTI. In this study, overall, 58.9 percent were with GSTM1 null genotype and 53.3 percent were with GSTT1 null genotype. There were a greater number of cases 63.3 percent with GSTM1 null genotype but the difference was not statistically significant. There was significant difference between the groups in GSTT1 but a greater number of controls were with null genotype (Table 31).

Some studies pointed to an association of GSTM1 null genotype with breast cancer risk in postmenopausal women (Helzlsouer *et al*., 1998; Charrier *et al*., 1999; Dunning *et al*., 1999) and in premenopausal women (Park *et al*., 2000). But most of the studies have found no association (Zhong *et al*., 1993; Ambrosone *et al*., 1995; Kelsey *et al*., 1997; Bailey *et al*., 1998; Ambrosone *et al*., 1999; Millikan *et al*., 2000). In most of the studies on GSTT1 genotype and breast cancer risk, no significant association has been found (Bailey *et al*., 1998; Helzlsouer *et al*., 1998; Curran *et al*., 2000; Millikan *et al*., 2000; Park *et al*., 2000). However, in one study the risk was found to be modified by the use of alcohol (Helzlsouer *et al*., 1998), and in another study a remarkably lower risk was suggested for premenopausal women lacking the GSTT1 gene (Garcia-Closas *et al*., 1999).

6.8. Association of Risk Factors with Mitochondrial DNA

Studies have shown the important role mitochondria play in both aging and cancer, and mutations in the mitochondrial genome have been associated with diverse forms of human disease. Defects in the mitochondrial genome and function are suspected to contribute to the development and progression of cancer (Ye *et al*., 2008). Strong evidence suggests that somatic mtDNA mutations play a role in breast carcinogenesis. The d-loop region has been shown to be a mutational 'hot spot' in human cancer. The incidence of somatic mutations in the d-loop region is found in all tumors examined to date and appears to be a universal feature of all cancers (Modica-Napolitano *et al*., 2007). The first study on breast cancer and association with mtDNA among the Mizos found mitochondrial gene alterations may attribute for risk (Ghatak *et al*., 2014). A depletion of COI subunit has been observed in breast cancer (Putignani *et al*., 2008). Several of the cancer-associated mutations found in mtDNA result in structural modifications of cytochrome c oxidase (Namslauer and Brzezinksi, 2009).

In this study, the mutation sites identified in d-loop with significant difference between the groups at 146 T>C, 16304 T>C, 16311 T>C and 16519 T>C were shown to have a predisposition to breast cancer by other studies (https://www.hmtvar.uniba.it/). The sites in CO1 5945 C>T and 6617 C>T were shown to be probably damaging; and 6392 T>C as a deleterious variant in Hmtvar as well (Table 39). These sites were further computed to check for any association with some of the risk factors. Some of the habits like use of betel quid and tobacco which did not show any significance in epidemiological analysis showed strong association for breast cancer with mutations T146C, A248Del, C5945T, T6392C, C6617T, T16304C and T16311C. Mutation T16519C showed association only with tobacco use. In vitro and in vivo experiments have shown that betel quid consumption can cause micronuclei and DNA adducts formation, chromosomal aberrations, allelic imbalances and sister chromatid exchange in oral mucosa cells (IARC, 2012). Calcium hydroxide a major content of slaked lime

in the presence of areca nut is responsible for the formation of ROS known to cause oxidative damage in the DNA of buccal mucosa cells of betel quid chewers (Nair *et al*., 2004). In one study, they found that buccal cells of healthy smokers harbored both a higher mtDNA mutation frequency and mutation density when compared with non-smokers (Tan *et al*., 2008). In another study, they found higher total number of genomic alterations in breast cancer tumors of betel quid chewers than to non-chewers (Kaushal *et al*., 2012).

There is no strong evidence linking the use of *tuibur* with breast cancer risk, but in this study, mutations C5945T (OR = 1.90, 95% CI: 1.43-2.53) and C6617T (OR = 1.50, 95% CI: 1.18-1.91) showed significant association. There are very few data linking cancer with the use of *tuibur*, one study found *tuibur* to increase stomach cancer (Malakar *et al*., 2012). Compared to smoking, higher levels of nicotine can enter systemic circulation from smokeless tobacco which indicates a much more potent effect through this route (Li *et al*., 2018). It generates ROS, oxidative stress, and associated DNA fragmentation in laboratory experiments. *Tuibur* exhibited significant toxicity by reducing the root growth of Allium bulbs and inducing tumor formation in the roots (Mahanta *et al*., 1998).

Mutations T152C, C5945T, C6617T, T16304C, T16311C and T16519C showed significant association with smoking. Mutation T16304C showed significant association with passive smoking. PAHs found in smoke have been shown to be mutagenic to breast cell lines, and as lipophilic compounds they are stored in adipose tissues (Li *et al*., 1996). Studies have found that women who inherited specific variants in genes involved in the metabolism of carcinogens found in tobacco might experience higher risks associated with smoking cigarettes (Morabia, 2002; Terry and Rohan, 2002; Slattery et al., 2008). Carcinogens found in tobacco are transported by the bloodstream, deposited and metabolically activated in the breast and surrounding

adipose tissues to the breast (Terry and Rohan, 2002). PAHs ingested or inhaled are converted to water-soluble derivatives mainly via oxidative activation by cytochrome P450 1A1 (CYP1A1) followed by detoxification by phase II enzymes such as GSTs.

Mutations T152C, A248Del, C5945T, T6392C, C6617T, T16304C, T16311C and T16519C showed significant association with disturbed sleep. Sleep deprivation induce cellular stress and oxidative damage to DNA. Studies have found that sleep deprivation induce cellular stress and oxidative damage to DNA. In one study, reduced sleep duration and sleep efficiency were associated with reduced mitochondrial DNA copy number in sleep duration discordant monozygotic twins whereby short sleep impairs health and longevity through mitochondrial stress (Wrede *et al*., 2015). Studies have also found that night shift work influence increase risk of hormone-related diseases including breast cancer because of sleep deprivation, circadian disruption, and exposure to light at night (Davis and Mirick, 2006).

Mutations C5945T, T6392C, C6617T, T16304C and T16519C showed significant association with miscarriage. Pregnancy is an inflammatory state exhibiting increased susceptibility to oxidative stress and increased DNA damage. There is strong evidence that lifestyle changes are increasing our risk of infertility and miscarriage, and that baseline DNA damage rises with age and couples in developed societies are delaying childbirth, placing them at further risk (Furness *et al*., 2011). More in depth studies is required to fully understand the association of mtDNA and miscarriage.

However, to fully understand the importance of mtDNA polymorphism and risk of cancer, it is necessary to identify genetic and epigenetic alteration in the nuclear DNA associated with mtDNA copy number changes; and to understand the role of synergy between different polymorphisms.

Conclusion

This study, in spite of many limitations, is the first in depth study of the epidemiology of breast cancer among the Mizos. The study is a retrospective study of 17 years, the earliest recorded case was in 1998 and no trace of records before that. There were 758 registered cases of which 9 were males. This is higher than the world data of 1 in 100. The number of sporadic cases to familial cases was 309 to 54 which was roughly 85:15 the ratio found worldwide. From 2003 – 2014 there was approximately 60 cases recorded in a year. The number of Hmar and Lusei tribes are the highest in both the case and control groups, from both the paternal and maternal side. Even though there are no records for population based on tribes, it seems that these two major tribes are having the highest population among the Mizos.

Since this is one of the few studies relating to epidemiology of breast cancer among the Mizos, interview was conducted with open-ended questions to allow the volunteers to answer freely without fear of being wrong. This allows the study to have a better understanding of many of the common regional risk factors which would have not been possible with just a 'yes' and 'no' response. The only difficulty is in categorizing the data especially on diet. When one could have the luxury of having meat every day, the other could have it only once a month. Affected individuals may associate their malignancy with foods perceived to be poor in nutritional value and over report them relative to unaffected controls. If a prospective study could also be conducted especially on dietary habits, stronger conclusions could have been made on some of the unique regional habits. Lifestyle appears to be a strong determinant of breast cancer risk and diet composition and nutritional status are important candidates. Understanding the role of diet in breast cancer is important because dietary factors are potentially modifiable risk factors on which preventive efforts may focus.

Further studies are required to identify all possible etiologic agents which was not covered in this study. It would be beneficial if further research could be conducted

on some of the risk factors identified in this study. The study would have been more significant if the age at which the subjects started their behavioural habits like smoking, use of betel quid, tobacco and tuibur could have been recorded. Exposures occurring at various stages of life, from as early as in utero up to age at diagnosis can potentially have an important impact on risk. Early life environment may contribute to breast cancer risk because the mammary tissue may be particularly susceptible to environmental influence at that time.

The limitation of this study is that a few of the results are influenced by recall bias and selection. But all respondents could provide reliable information, especially breast cancer volunteers which were cross-checked with their medical records. The other limitation is interviewing time. The interview time was fixed at 30-40 minutes, care was taken not to take too much time to not overwhelm or make the volunteers lose interest with too many questions.

CHAPTER 7

SUMMARY

7. Summary

- Incidence is more closely associated to age than to any other risk factor, it increases rapidly during the reproductive years and then more slowly after 50 years of age in this study. The highest incidence was observed in ages between 40 and 49 years.
- Strong association with risk was observed for smoking, frequent use of fermented pork fat, having hypertension, late age at menopause and at first childbirth.
- Among the premenopausal groups, heavy physical activity was observed to increase risk. A person can be physically active and yet spend a substantial amount of time being sedentary. If further study could be conducted for association of physical activity in detail, the results might be beneficial. Among the postmenopausal groups, association was observed with being overweight or obese, smoking and frequent use of sodium bicarbonate.
- A follow-up after 5 years showed that survival was influenced by three important risk factors - they had better survival if they do not smoke, had a lifestyle with moderate physical activity and shorter lifetime cumulative number of menstrual cycles.
- Association of Glutathione s transferase (M1 & T1) with risk factors does not have any relevant significance in this study group.
- Significant mtDNA mutations in this study either had predisposition to breast cancer or probably damaging or deleterious variant as established from other published data. Interaction of these mutations with behavioural habits reveal significant association to increase risk with use of betel quid, tobacco, tuibur, smoking, reduced sleep duration and sleep efficiency. Interaction of miscarriages and older age of mother with these mutations also had significant association.
- The established factors that are modifiable observed in this study are late age at first child birth and being obese or overweight after menopause; the non-modifiable are age and late age at menopause; the suspected are smoking cigarette and hypertension; the unclear but suspected found specific to this region is frequent use of soda bicarbonate and fermented pork fat.

VIII. APPENDIX

1. List of Acronyms

Abbreviated	**Full Form**
°	Degree
%	Percent
C	Celsius
hr	Hour
min	Minutes
sec	Seconds
kg	Kilograms
g	Gram
mg	Milli Gram
μg	Micro Gram
L	Litre
ml	Milli Litre
μL	Micro Litre
m	Meter
cm	Centimeter
mM	Milli Molar
nM	Nano Molar
bp	Base Pair
np	Nucleotide Position
rpm	Revolutions per minute
AAR	Age Adjusted Rate
AC	Anno Domini
AICR	American Institute for Cancer Research
ATM	Ataxia Telangiectasia Mutation
ATPase	ATP Synthase
BARD1	BRCA1 associated RING domain protein 1
BC	Before Christ
BMI	Body mass index
BPA	Bisphenol A
BRCA	Breast Cancer gene
BSA	Body Surface Area
Ca	Cancer Cases
CDH1	Cadherin1
CHEK2	Checkpoint kinase 2
CI	Confidence Interval
CIS	Carcinoma In situ
Co	Control
CO1	Cytochrome C Oxidase 1
CR	Crude Ratio
Cyt b	Cytochrome b
DCIS	Ductal Carcinoma In situ
DDT	Dichloro-diphenyl-trichloroethane

D-Loop	Displacement Loop
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic acid
ELF	Extremely Low Frequency
EMF	Electro Magnetic Field
EMR	Electromagnetic Radiation
ER	Estrogen Receptor
ESR1	Estrogen Receptor 1
FCB	First Child at Birth
GLOBOCAN	Global Cancer Observatory
GSH	Glutathione
GST	Glutathione s-transferase
GSTM1	Glutathione s-transferase mu 1
GSTT1	Glutathione s-transferase theta 1
HAA	Heterocyclic Aromatic Amines
HER2	Human Epidermal Growth Factor Receptor 2
HRT	Hormone Replacement Therapy
IARC	International Agency for Research on Cancer
IDC	Invasive Ductal Carcinoma
ILC	Invasive Lobular Carcinoma
LCB	Last Child at Birth
LCIS	Lobular Carcinoma In situ
MET	Metabolic Equivalent of Task
mtDNA	Mitochondrial DNA
NBS	Nijmegen breakage syndrome
NCRP	National Cancer Registry Program
ND	NADH Dehydrogenase
OR	Odds Ratio
PAH	Polycyclic Aromatic Hydrocarbons
PALB2	Partner and Localizer of BRCA2
PBCR	Population Based Cancer Registry
PBS	Phosphate-buffered saline
PR	Progesterone Receptor
PTEN	Phosphatase and tensin homolog
RAD51C	Restriction site associated DNA Homolog 51 C
RAD51D	Restriction site associated DNA Homolog 51 D
ROS	Reactive Oxygen Species
SD	Standard Deviation
STK11	Serine/Threonine Kinase 11
TNM	Tumor, Node and Metastasis
TP53	Tumor Protein p53
WCRF	World Cancer Research Fund
WHO	World Health Organization
WHR	Waist to Hip Ratio
XRCC	X-Ray Repair Cross Complementing

2. Ethics Approval Certificate

IEC, CIVIL HOSPITAL, AIZAWL.
COMMUNICATION TO THE PRINCIPAL INVESTIGATOR/GUIDE BY THE MEMBER SECRETARY, INSTITUTIONAL ETHICS COMMITTEE

No.B.12018/1/13-CH(A)/IEC/28 Dated: 15th of October, 2014

To,

Dr. G. Gurusubramanian,
Associate Prof., Department of Zoology,
Mizoram University.

The Institutional Ethics Committee in its meeting held on 14th of October 2014, has reviewed and discussed your application to conduct the clinical trial/project entitled; **Mutational Analysis of Mitochondrial and Glutathione S-Transferase Genes Associated with Breast Cancer Risk Among Mizo Population.**

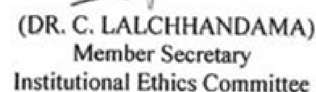
Sponsored by : **National Institute of Biomedical Genomics, West Bengal and Department of Biotechnology, New Delhi.**

Code no. : ________________________

The following documents were reviewed:

a. Trial Protocol (including protocol amendments)/project, dated 4th of August, 2014.
b. Investigator's Brochure, dated 4th of August, 2014.
c. Patient Information Sheet and Informed Consent Form (including updates if any) in Hindi, English and/or vernacular language.
d. Proposed methods for patient accrual including advertisement (s) etc. proposed to be used for the purpose.
e. Current CV of investigator from outside Civil Hospital, Aizawl.
f. Insurance Policy/Compensation for participation and for serious adverse events occurring during the study participation.
g. Investigator's Agreement with the Sponsor.
h. Investigator's Undertaking.
i. Ethics Committee Proforma.
j. DCGI approval letter/submission letter, if any.
k. Civil Hospital. Aizawl Case Report Form
l. Any other/additional documents

Decision of Committee: Approved/~~Not Approved~~



(DR. C. LALCHHANDAMA)
Member Secretary
Institutional Ethics Committee

Secretary
Institutional Ethical Committee
Civil Hospital, Aizawl

3. Consent

TOPIC: Mutational analysis of mitochondrial and glutathione s-tranferase genes associated with breast cancer risk among Mizo population

HMING/ID : ____

CONTACT : ____

1. He zirchianna hi hrilhfiahna hun siamsak ka ni.
2. He zirchiannaah hian keima remtihna ngeiin ka tel. Ka duh hun hunah, chhan pawh sawi loin, ka inenkawlna leh ka dikna engmah nghawng nei loin ka inhnukdawk thei tih ka hria.
3. He zirchianna atan leh thildang hemi kaihhnawih chi reng rengah zirchiangtute, ethics committee-te leh thuneitu-tein ka remtihna tel kher loin (ka ban hnuah pawh) ka hriselna lam record an en thei tih ka hria a. Mahse ka nihna chu engti kawng mahin puanzar a ni lo ang tih ka hre bawk.
4. Ka biological sample te hi he zirchianna (research) atan hian pek ka remti thlap a, ka hriselna khawih pawi thei engmah a awm lo tih ka hria e.
5. He zirchianna aṭanga hmuhchhuahte hi Science lam thil bik atan chuan duh ang anga hman ka remti thlap e.
6. He zirchianna aṭang hian hlawkna engmah beisei ka nei lo. Nakin hun, midang tan hlawkna tur a ni tih ka hrebawk.
7. He zirchianna-ah hian tel ka rem ti e.

DATE: SIGNATURE

4. Questionnaire (Mizo)

TOPIC: Mutational analysis of mitochondrial and glutathione s-transferase genes associated with breast cancer risk among Mizo population

ID No.: ____________________________ Date: __________________

1) 1) Hming ____

2) Pian ni ____

3) Zirna Zir lo____ Primary/Middle_____ High School/12_______ Graduate____

4) Kum 5 chin chenna _________________ 5) Ei zawnna _____________________

6) San zawng, rih zawng ________Feet _______Kg 7) BMI __________

8) Taksa sawizawi Ngailo______zeuh zeuh_______awm tawk vel_____ nasa_______

9) Tihṭhin
   a) Sahdah ________ Khaini________ Tuibur________ Meizial _________
      Meizial ni 1-a zuk zat__________ Mei zu Chenpui________________
   b) Zu In ngai miah lo ________ In ve zeuh zeuh_________Nitin In ________
      Ni khat-ah________Kum__________Eng chi________________________
   c) Kuhva Ei ngai lo _______ zeuh zeuh _____ Ei______________________
   d) Ni tin ei ṭhinah sa-um _______ soda/bai ______ thei ______ thlai_______
   e) Sa i ei ṭhin em? Eng sa nge i ei ngun bik deuh? Aw ______ Aih _____
      Ar___ Sangha___ Vawksa___B awngsa____ Kelsa____ Sa Dang_______
      Karkhatah ei zah _________ emaw Thla khat________________
   f) Tui ni khatah Litre 1 tlinglo ______Litre 1_____Litre 2 _____ Litre 3_____
   g) Chi ei nasa lo_______ ei nasa ve tho________ ei nasa __________
   h) Darkar eng zah nge mut tlangpui________Mu ṭha lo _______Mu ṭha _____

10) Have you been diagnosed with Breast Cancer? (Data-sheet-ah awm loh chuan)

   a) No

   b) Yes (left______/right_______/both______) Kum eng zah nihin__________

   c) Type – IDC____ DCIS_____ ILC_____ LCIS_____ Adang___________

11) Eng vanga doctor pan nge? Tui-ril/thisen chhuak, bawk, hnute na

Tui-ril chhuak_______ Thisen chuak______ Hnute na _____ A hmur tlum_____

A vun sen____________ Bawk ___________________________________

Engtia hmu nge? _________________ Doctor pan nghal em? _____________

12) Hormone Replacement Therapy & Oral contraceptive:

a) Period nei ṭha lo______ ṭha ________Neihin Na/buai em _____________

b) Indanna ei ngai_________________________ Ngai lo_______________

c) Thi Hul lo_________ Hul_______ Kum eng zahah nge HUL__________

d) Hul dawnin buaina Awm Lo______ Awm__________________________

Engtia enkawl nge_____________Damdawi ei__________ Enkawl Lo____

13) Pesticide exposure/Electromagnetic Fields:

a) Huan thlai i enkawl em?________ Hlo hmang Lo _______ Hmang ________

b) In in bul hnai velah Phone tower____ Electric ban______ Transformer_____

14) Reproductive History -

a) Kum eng zatah nge thi i neih ṭan_________ pasal neih ________________

b) Fa nei lo______neih zat _______Chhiat awm_______Chhiat awm lo_______

c) Kum eng zah i nihin nge fa hmasa ber i neih, a hnuhnung ber kum eng zatah nge i neih FCB_____________________ LCB _________________________

d) Fate hnute tui Hnektir vek ______ Hnektir vek lo ______ hnektir lo ______

Hnektir chhung Thla 6 tlinglo______Thla 6-kum _____ Kum 1 aia tam ____

15) Medical History/Natna dang i nei em ___________________________________

16) Chhungte cancer vei an awm em? Inchhun dan leh eng cancer nge ___________

____

17) Unau zat _____ Eng zahna_______ Hnute hne _____ Piana nu kum zat _______

CHAPTER 9

REFERENCES

9. References

ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (2002). Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *JAMA*, *288*(23), 2981–2997. https://doi.org/10.1001/jama.288.23.2981

Ambrosone, C. B., Coles, B. F., Freudenheim, J. L., & Shields, P. G. (1999). Glutathione-S-transferase (GSTM1) genetic polymorphisms do not affect human breast cancer risk, regardless of dietary antioxidants. *The Journal of nutrition*, *129*(2S Suppl), 565S–568S. https://doi.org/10.1093/jn/129.2.565S

Ambrosone, C. B., Freudenheim, J. L., Graham, S., Marshall, J. R., Vena, J. E., Brasure, J. R., Laughlin, R., Nemoto, T., Michalek, A. M., & Harrington, A. (1995). Cytochrome P4501A1 and glutathione S-transferase (M1) genetic polymorphisms and postmenopausal breast cancer risk. *Cancer research*, *55*(16), 3483–3485.

Ambrosone, C.B., & Shields, P.G. (1999). Smoking as a risk factor for breast cancer. In Bowcock, A.M. (ed.) *Breast Cancer: Molecular Genetics, Pathogenesis and Therapeutics,* 519-36.

Andrieu, N., Prevost, T., Rohan, T. E., Luporsi, E., Lê, M. G., Gerber, M., Zaridze, D. G., Lifanova, Y., Renaud, R., Lee, H. P., & Duffy, S. W. (2000). Variation in the interaction between familial and reproductive factors on the risk of breast cancer according to age, menopausal status, and degree of familiality. *International journal of epidemiology*, *29*(2), 214–223. https://doi.org/10.1093/ije/29.2.214

Andrieu, N., Smith, T., Duffy, S., Zaridze, D. G., Renaud, R., Rohan, T., Gerber, M., Luporsi, E., Lê, M., Lee, H. P., Lifanova, Y., & Day, N. E. (1998). The effects of interaction between familial and reproductive factors on breast cancer risk: a combined analysis of seven case-control studies. *British journal of cancer*, *77*(9), 1525–1536. https://doi.org/10.1038/bjc.1998.251

Apter, D., Reinilä, M., & Vihko, R. (1989). Some endocrine characteristics of early menarche, a risk factor for breast cancer, are preserved into adulthood. *International journal of cancer*, *44*(5), 783–787. https://doi.org/10.1002/ijc.2910440506

Arand, M., Mühlbauer, R., Hengstler, J., Jäger, E., Fuchs, J., Winkler, L., & Oesch, F. (1996). A multiplex polymerase chain reaction protocol for the simultaneous analysis of the glutathione S-transferase GSTM1 and GSTT1 polymorphisms. *Analytical biochemistry*, *236*(1), 184–186. https://doi.org/10.1006/abio.1996.0153

Ariel, I.M. (1987). Breast cancer, a historic review: is the past prologue? In: Ariel IM, Cleary JB, Eds. *Breast Cancer Diagnosis and Treatment*. New York: McGraw-Hill 3-26

Armstrong, B., & Doll, R. (1975). Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *International journal of cancer*, *15*(4), 617–631. https://doi.org/10.1002/ijc.2910150411

Atkinson, C., Lampe, J. W., Tworoger, S. S., Ulrich, C. M., Bowen, D., Irwin, M. L., Schwartz, R. S., Rajan, B. K., Yasui, Y., Potter, J. D., & McTiernan, A. (2004). Effects of a moderate intensity exercise intervention on estrogen metabolism in postmenopausal women. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer*

Research, cosponsored by the American Society of Preventive Oncology, *13*(5), 868–874.

Babita, K. N., Singh, M., Malik, J. S., & Kalhan, M. (2014). Breastfeeding reduces breast cancer risk: a case-control study in north India. *International journal of preventive medicine*, *5*(6), 791–795.

Bailey, L. R., Roodi, N., Verrier, C. S., Yee, C. J., Dupont, W. D., & Parl, F. F. (1998). Breast cancer and CYPIA1, GSTM1, and GSTT1 polymorphisms: evidence of a lack of association in Caucasians and African Americans. *Cancer research*, *58*(1), 65–70.

Ballard-Barbash, R. (1994). Anthropometry and breast cancer. Body size--a moving target. *Cancer*, *74*(3 Suppl), 1090–1100. https://doi.org/10.1002/1097-0142(19940801)74:3+<1090::aid-cncr2820741518>3.0.co;2-x

Barekati, Z., Radpour, R., Kohler, C., Zhang, B., Toniolo, P., Lenner, P., Lv, Q., Zheng, H., & Zhong, X. Y. (2010). Methylation profile of TP53 regulatory pathway and mtDNA alterations in breast cancer patients lacking TP53 mutations. *Human molecular genetics*, *19*(15), 2936–2946. https://doi.org/10.1093/hmg/ddq199

Barnes, S. L., Singletary, K. W., & Frey, R. (2000). Ethanol and acetaldehyde enhance benzo[a]pyrene-DNA adduct formation in human mammary epithelial cells. *Carcinogenesis*, *21*(11), 2123–2128. https://doi.org/10.1093/carcin/21.11.2123

Bartsch, H., Nair, J., & Owen, R. W. (1999). Dietary polyunsaturated fatty acids and cancers of the breast and colorectum: emerging evidence for their role as risk modifiers. *Carcinogenesis*, *20*(12), 2209–2218. https://doi.org/10.1093/carcin/20.12.2209

Bartsch, H., Nair, U., Risch, A., Rojas, M., Wikman, H., & Alexandrov, K. (2000). Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *9*(1), 3–28.

Bauer, K. R., Brown, M., Cress, R. D., Parise, C. A., & Caggiano, V. (2007). Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer*, *109*(9), 1721–1728. https://doi.org/10.1002/cncr.22618

Benson, J. R., Weaver, D. L., Mittra, I., & Hayashi, M. (2003). The TNM staging system and breast cancer. *The Lancet. Oncology*, *4*(1), 56–60. https://doi.org/10.1016/s1470-2045(03)00961-6

Berkey, C. S., Frazier, A. L., Gardner, J. D., & Colditz, G. A. (1999). Adolescence and breast carcinoma risk. *Cancer*, *85*(11), 2400–2409. https://doi.org/10.1002/(sici)1097-0142(19990601)85:11<2400::aid-cncr15>3.0.co;2-o

Bernstein, L., Henderson, B. E., Hanisch, R., Sullivan-Halley, J., & Ross, R. K. (1994). Physical exercise and reduced risk of breast cancer in young women. *Journal of the National Cancer Institute*, *86*(18), 1403–1408. https://doi.org/10.1093/jnci/86.18.1403

Bernstein, L., Ross, R. K., Lobo, R. A., Hanisch, R., Krailo, M. D., & Henderson, B. E. (1987). The effects of moderate physical activity on menstrual cycle patterns in adolescence: implications for breast cancer prevention. *British journal of cancer*, *55*(6), 681–685. https://doi.org/10.1038/bjc.1987.139

Bianchi, N. O., Bianchi, M. S., & Richard, S. M. (2001). Mitochondrial genome instability in human cancers. *Mutation research*, *488*(1), 9–23. https://doi.org/10.1016/s1383-5742(00)00063-6

Black, S. M., Beggs, J. D., Hayes, J. D., Bartoszek, A., Muramatsu, M., Sakai, M., & Wolf, C. R. (1990). Expression of human glutathione S-transferases in Saccharomyces cerevisiae confers resistance to the anticancer drugs adriamycin and chlorambucil. *The Biochemical journal*, *268*(2), 309–315. https://doi.org/10.1042/bj2680309

Blackburn, A. C., Tzeng, H. F., Anders, M. W., & Board, P. G. (2000). Discovery of a functional polymorphism in human glutathione transferase zeta by expressed sequence tag database analysis. *Pharmacogenetics*, *10*(1), 49–57. https://doi.org/10.1097/00008571-200002000-00007

Board, P. G., Baker, R. T., Chelvanayagam, G., & Jermiin, L. S. (1997). Zeta, a novel class of glutathione transferases in a range of species from plants to humans. *The Biochemical journal*, *328 (Pt 3)*(Pt 3), 929–935. https://doi.org/10.1042/bj3280929

Bonner, M. R., Han, D., Nie, J., Rogerson, P., Vena, J. E., Muti, P., Trevisan, M., Edge, S. B., & Freudenheim, J. L. (2005). Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *14*(1), 53–60.

Bouker, K. B., & Hilakivi-Clarke, L. (2000). Genistein: does it prevent or promote breast cancer? *Environmental health perspectives*, *108*(8), 701–708. https://doi.org/10.1289/ehp.00108701

Breasted, J.H. Ed. (1930). The Edwin Smith Surgical papyrus. Chicago, Illinois: The University Chicago Press special edition 1984.

Brinton, L. A., Schairer, C., Hoover, R. N., & Fraumeni, J. F., Jr (1988). Menstrual factors and risk of breast cancer. *Cancer investigation*, *6*(3), 245–254. https://doi.org/10.3109/07357908809080645

Broca, P. (1866). *Traite des tumeurs*

Brody, J. G., Moysich, K. B., Humblet, O., Attfield, K. R., Beehler, G. P., & Rudel, R. A. (2007). Environmental pollutants and breast cancer: epidemiologic studies. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, *109*, 2667-2711.

Brunnemann, K. D., & Hoffmann, D. (1992). Chemical Composition of Smokeless Tobacco Or Health*: An International Perspective*, (92), 96.

Bruzzi, P., Negri, E., La Vecchia, C., Decarli, A., Palli, D., Parazzini, F., & Del Turco, M. R. (1988). Short term increase in risk of breast cancer after full term pregnancy. *BMJ (Clinical research ed.)*, *297*(6656), 1096–1098. https://doi.org/10.1136/bmj.297.6656.1096

Butt, S., Borgquist, S., Anagnostaki, L., Landberg, G., & Manjer, J. (2014). Breastfeeding in relation to risk of different breast cancer characteristics. *BMC research notes*, *7*, 216. https://doi.org/10.1186/1756-0500-7-216

Campbell, K. L., Winters-Stone, K. M., Wiskemann, J., May, A. M., Schwartz, A. L., Courneya, K. S., Zucker, D. S., Matthews, C. E., Ligibel, J. A., Gerber, L. H., Morris, G. S., Patel, A. V., Hue, T. F., Perna, F. M., & Schmitz, K. H. (2019). Exercise Guidelines for Cancer Survivors: Consensus Statement from International Multidisciplinary Roundtable. *Medicine and science in sports and exercise*, *51*(11), 2375–2390. https://doi.org/10.1249/MSS.0000000000002116

Castelló, A., Martín, M., Ruiz, A., Casas, A. M., Baena-Cañada, J. M., Lope, V., Antolín, S., Sánchez, P., Ramos, M., Antón, A., Muñoz, M., Bermejo, B., De Juan-Ferré, A., Jara, C., Chacón, J. I., Jimeno, M. A., Rosado, P., Díaz, E., Guillem, V., Lluch, A., … EpiGEICAM Researchers (2015). Lower Breast Cancer Risk among Women following the World Cancer Research Fund and American Institute for Cancer Research Lifestyle Recommendations: EpiGEICAM Case-Control Study. *PloS one*, *10*(5), e0126096. https://doi.org/10.1371/journal.pone.0126096

Chang, K. J., Lee, T. T., Linares-Cruz, G., Fournier, S., & de Ligniéres, B. (1995). Influences of percutaneous administration of estradiol and progesterone on human breast epithelial cell cycle in vivo. *Fertility and sterility*, *63*(4), 785–791.

Charrier, J., Maugard, C. M., Le Mevel, B., & Bignon, Y. J. (1999). Allelotype influence at glutathione S-transferase M1 locus on breast cancer susceptibility. *British journal of cancer*, *79*(2), 346–353. https://doi.org/10.1038/sj.bjc.6690055

Chawngkunga, C. (1996). Genealogical tree of Mizos. Mizoram Publication Board, Art & Culture, Government of Mizoram. 1-347.

Chen, C., Ma, X., Zhong, M., & Yu, Z. (2010). Extremely low-frequency electromagnetic fields exposure and female breast cancer risk: a meta-analysis based on 24,338 cases and 60,628 controls. *Breast cancer research and treatment*, *123*, 569-576.

Chie, W. C., Hsieh, C., Newcomb, P. A., Longnecker, M. P., Mittendorf, R., Greenberg, E. R., Clapp, R. W., Burke, K. P., Titus-Ernstoff, L., Trentham-Dietz, A., & MacMahon, B. (2000). Age at any full-term pregnancy and breast cancer risk. *American journal of epidemiology*, *151*(7), 715–722. https://doi.org/10.1093/oxfordjournals.aje.a010266

Chlebowski, R. T. (2013). Nutrition and physical activity influence on breast cancer incidence and outcome. *Breast (Edinburgh, Scotland)*, *22 Suppl 2*, S30–S37. https://doi.org/10.1016/j.breast.2013.07.006

Chlebowski, R. T., Hendrix, S. L., Langer, R. D., Stefanick, M. L., Gass, M., Lane, D., Rodabough, R. J., Gilligan, M. A., Cyr, M. G., Thomson, C. A., Khandekar, J., Petrovitch, H., McTiernan, A., & WHI Investigators (2003). Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial. *JAMA*, *289*(24), 3243–3253. https://doi.org/10.1001/jama.289.24.3243

Chu, S. Y., Lee, N. C., Wingo, P. A., Senie, R. T., Greenberg, R. S., & Peterson, H. B. (1991). The relationship between body mass and breast cancer among women enrolled in the Cancer and Steroid Hormone Study. *Journal of clinical epidemiology*, *44*(11), 1197–1206. https://doi.org/10.1016/0895-4356(91)90152-y

Claus, E. B., Schildkraut, J. M., Thompson, W. D., & Risch, N. J. (1996). The genetic attributable risk of breast and ovarian cancer. *Cancer*, *77*(11), 2318–2324. https://doi.org/10.1002/(SICI)1097-0142(19960601)77:11<2318::AID-CNCR21>3.0.CO;2-Z

Claus, E. B., Schildkraut, J., Iversen, E. S., Jr, Berry, D., & Parmigiani, G. (1998). Effect of BRCA1 and BRCA2 on the association between breast cancer risk and family history. *Journal of the National Cancer Institute*, *90*(23), 1824–1829. https://doi.org/10.1093/jnci/90.23.1824

Clavel-Chapelon, F., & E3N Group (2002). Cumulative number of menstrual cycles and breast cancer risk: results from the E3N cohort study of French women. *Cancer causes & control: CCC*, *13*(9), 831–838. https://doi.org/10.1023/a:1020684821837

Clavel-Chapelon, F., & Gerber, M. (2002). Reproductive factors and breast cancer risk. Do they differ according to age at diagnosis?. *Breast cancer research and treatment*, *72*(2), 107–115. https://doi.org/10.1023/a:1014891216621

Cleary, M. P., & Grossmann, M. E. (2009). Minireview: Obesity and breast cancer: the estrogen connection. *Endocrinology*, *150*(6), 2537–2542. https://doi.org/10.1210/en.2009-0070

Colditz, G. A., Stampfer, M. J., Willett, W. C., Hennekens, C. H., Rosner, B., & Speizer, F. E. (1990). Prospective study of estrogen replacement therapy and risk of breast cancer in postmenopausal women. *JAMA*, *264*(20), 2648–2653.

Collaborative Group on Hormonal Factors in Breast Cancer (1996). Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. *Lancet (London, England)*, *347*(9017), 1713–1727. https://doi.org/10.1016/s0140-6736(96)90806-5

Collaborative Group on Hormonal Factors in Breast Cancer (1997). Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet (London, England)*, 350(9084), 1047–1059.

Collaborative Group on Hormonal Factors in Breast Cancer (2002). Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet (London, England)*, *360*(9328), 187–195. https://doi.org/10.1016/S0140-6736(02)09454-0

Combes, B., & Stakelum, G. S. (1961). A liver enzyme that conjugates sulfobromophthalein sodium with glutathione. *The Journal of clinical investigation*, *40*(6), 981–988. https://doi.org/10.1172/JCI104337

Connor, K., Ramamoorthy, K., Moore, M., Mustain, M., Chen, I., Safe, S., Zacharewski, T., Gillesby, B., Joyeux, A., & Balaguer, P. (1997). Hydroxylated polychlorinated biphenyls (PCBs) as estrogens and antiestrogens: structure-activity relationships. *Toxicology and applied pharmacology*, *145*(1), 111–123. https://doi.org/10.1006/taap.1997.8169

Cook, J.W., Hewett, C.L. & Hieger, I. (1933). The isolation of a cancer producing hydrocarbon from coal tar. *Journal of the Chemical Society*, 395-405.

Coughlin, S. S., Calle, E. E., Teras, L. R., Petrelli, J., & Thun, M. J. (2004). Diabetes mellitus as a predictor of cancer mortality in a large cohort of US adults. *American journal of epidemiology*, *159*(12), 1160–1167. https://doi.org/10.1093/aje/kwh161

Covarrubias, D., Bai, R. K., Wong, L. C., & Leal, S. M. (2008). Mitochondrial DNA variant interactions modify breast cancer risk. *Journal of human genetics*, *53*(10), 924–928. https://doi.org/10.1007/s10038-008-0331-x

Cserni, G., Chmielik, E., Cserni, B., & Tot, T. (2018). The new TNM-based staging of breast cancer. *Virchows Archiv : an international journal of pathology*, *472*(5), 697–703. https://doi.org/10.1007/s00428-018-2301-9

Cui, Y., Miller, A. B., & Rohan, T. E. (2006). Cigarette smoking and breast cancer risk: update of a prospective cohort study. *Breast cancer research and treatment*, *100*(3), 293–299. https://doi.org/10.1007/s10549-006-9255-3

Curran, J. E., Weinstein, S. R., & Griffiths, L. R. (2000). Polymorphisms of glutathione S-transferase genes (GSTM1, GSTP1 and GSTT1) and breast cancer susceptibility. *Cancer letters*, *153*(1-2), 113–120. https://doi.org/10.1016/s0304-3835(00)00361-x

Czeisler, C. A., & Klerman, E. B. (1999). Circadian and sleep-dependent regulation of hormone release in humans. *Recent progress in hormone research*, *54*, 97–132.

Dahlgren, A., Lundmark, P., Axelsson, T., Lind, L., & Syvänen, A. C. (2008). Association of the estrogen receptor 1 (ESR1) gene with body height in adult males from two Swedish population cohorts. *PloS one*, *3*(3), e1807. https://doi.org/10.1371/journal.pone.0001807

Das, S. K. (2003). Harmful health effects of cigarette smoking. *Molecular and cellular biochemistry*, *253*(1-2), 159–165. https://doi.org/10.1023/a:1026024829294

Davis, D. L., Bradlow, H. L., Wolff, M., Woodruff, T., Hoel, D. G., & Anton-Culver, H. (1993). Medical hypothesis: xenoestrogens as preventable causes of breast cancer. *Environmental health perspectives*, *101*(5), 372–377. https://doi.org/10.1289/ehp.93101372

Davis, S., & Mirick, D. K. (2006). Circadian disruption, shift work and the risk of cancer: a summary of the evidence and studies in Seattle. *Cancer causes & control : CCC*, *17*(4), 539–545. https://doi.org/10.1007/s10552-005-9010-9

Davis, S., Mirick, D. K., & Stevens, R. G. (2001). Night shift work, light at night, and risk of breast cancer. *Journal of the National Cancer Institute*, *93*(20), 1557–1562. https://doi.org/10.1093/jnci/93.20.1557

Davis, S., Mirick, D. K., & Stevens, R. G. (2002). Residential magnetic fields and the risk of breast cancer. *American journal of epidemiology*, *155*(5), 446–454. https://doi.org/10.1093/aje/155.5.446

Davis, S., Mirick, D. K., Chen, C., & Stanczyk, F. Z. (2006). Effects of 60-Hz magnetic field exposure on nocturnal 6-sulfatoxymelatonin, estrogens, luteinizing hormone, and follicle-stimulating hormone in healthy reproductive-age women: results of a crossover trial. *Annals of epidemiology*, *16*(8), 622–631. https://doi.org/10.1016/j.annepidem.2005.11.005

De Crée, C., Van Kranenburg, G., Geurten, P., Fujimori, Y., & Keizer, H. A. (1997). 4-Hydroxycatecholestrogen metabolism responses to exercise and training: possible implications for menstrual cycle irregularities and breast cancer. *Fertility and sterility*, *67*(3), 505–516. https://doi.org/10.1016/s0015-0282(97)80078-7

De Moulin, D. (1983). A short history of breast cancer. Boston: Martinus Nijhoff, 1-107.

Deandrea, S., Talamini, R., Foschi, R., Montella, M., Dal Maso, L., Falcini, F., La Vecchia, C., Franceschi, S., & Negri, E. (2008). Alcohol and breast cancer risk defined by estrogen and progesterone receptor status: a case-control study. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *17*(8), 2025–2028. https://doi.org/10.1158/1055-9965.EPI-08-0157

Decarli, A., La Vecchia, C., Negri, E., & Franceschi, S. (1996). Age at any birth and breast cancer in Italy. *International journal of cancer*, *67*(2), 187–189. https://doi.org/10.1002/(SICI)1097-0215(19960717)67:2<187::AID-IJC6>3.0.CO;2-P

Dees, C., Askari, M., Foster, J. S., Ahamed, S., & Wimalasena, J. (1997). DDT mimicks estradiol stimulation of breast cancer cells to enter the cell cycle. *Molecular carcinogenesis*, *18*(2), 107–114.

Dees, C., Foster, J. S., Ahamed, S., & Wimalasena, J. (1997). Dietary estrogens stimulate human breast cells to enter the cell cycle. *Environmental health perspectives*, *105 Suppl 3*(Suppl 3), 633–636. https://doi.org/10.1289/ehp.97105s3633

Deurenberg, P., Weststrate, J. A., & Seidell, J. C. (1991). Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. *The British journal of nutrition*, *65*(2), 105–114. https://doi.org/10.1079/bjn19910073

Diamandopoulos, G. T. (1996). Cancer: an historical perspective. *Anticancer research*, *16*(4A), 1595–1602.

Dorgan, J. F., Baer, D. J., Albert, P. S., Judd, J. T., Brown, E. D., Corle, D. K., Campbell, W. S., Hartman, T. J., Tejpar, A. A., Clevidence, B. A., Giffen, C. A., Chandler, D. W., Stanczyk, F. Z., & Taylor, P. R. (2001). Serum hormones and the alcohol-breast cancer association in postmenopausal women. *Journal of the National Cancer Institute*, *93*(9), 710–715. https://doi.org/10.1093/jnci/93.9.710

Dorgan, J. F., Brown, C., Barrett, M., Splansky, G. L., Kreger, B. E., D'Agostino, R. B., Albanes, D., & Schatzkin, A. (1994). Physical activity and risk of breast cancer in the Framingham Heart Study. *American journal of epidemiology*, *139*(7), 662–669. https://doi.org/10.1093/oxfordjournals.aje.a117056

Dumitrescu, R. G., & Shields, P. G. (2005). The etiology of alcohol-induced breast cancer. *Alcohol (Fayetteville, N.Y.)*, *35*(3), 213–225. https://doi.org/10.1016/j.alcohol.2005.04.005

Dunning, A. M., Healey, C. S., Pharoah, P. D., Teare, M. D., Ponder, B. A., & Easton, D. F. (1999). A systematic review of genetic polymorphisms and breast cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *8*(10), 843–854.

Durando, M., Kass, L., Piva, J., Sonnenschein, C., Soto, A. M., Luque, E. H., & Muñoz-de-Toro, M. (2007). Prenatal bisphenol A exposure induces preneoplastic lesions in the mammary gland in Wistar rats. *Environmental health perspectives*, *115*(1), 80–86. https://doi.org/10.1289/ehp.9282

Easton, D. F. (2002). Familial risks of breast cancer. *Breast cancer research : BCR*, *4*(5), 179–181. https://doi.org/10.1186/bcr448

Easton, D., Ford, D., & Peto, J. (1993). Inherited susceptibility to breast cancer. *Cancer surveys*, *18*, 95–113.

Edgren, R. A. (1991). Oral contraception: a review. *International journal of fertility*, *36 Suppl 3*, 16–25.

Edler, D., Klein, J., Antonelli, A., & Silvestro, D. (2021). raxmlGUI 2.0: a graphical interface and toolkit for phylogenetic analyses using RAxML. Methods in Ecology and Evolution, 12(2), 373-377. https://doi.org/10.1111/2041-210X.13512

Egan, K. M., Cai, Q., Shu, X. O., Jin, F., Zhu, T. L., Dai, Q., Gao, Y. T., & Zheng, W. (2004). Genetic polymorphisms in GSTM1, GSTP1, and GSTT1 and the risk for breast cancer: results from the Shanghai Breast Cancer Study and meta-analysis. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *13*(2), 197–204. https://doi.org/10.1158/1055-9965.epi-03-0294

Ellison, P. T., & Lager, C. (1986). Moderate recreational running is associated with lowered salivary progesterone profiles in women. *American journal of obstetrics and gynecology*, *154*(5), 1000–1003. https://doi.org/10.1016/0002-9378(86)90737-4

Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, *19*(5), 403–410. https://doi.org/10.1111/j.1365-2559.1991.tb00229.x

Emerit, I. (1994). Reactive oxygen species, chromosome mutation, and cancer: possible role of clastogenic factors in carcinogenesis. *Free radical biology & medicine*, *16*(1), 99–109. https://doi.org/10.1016/0891-5849(94)90246-1

Emons, J., Chagin, A. S., Malmlöf, T., Lekman, M., Tivesten, A., Ohlsson, C., Wit, J. M., Karperien, M., & Sävendahl, L. (2010). Expression of vascular endothelial growth factor in the growth plate is stimulated by estradiol and increases during pubertal development. *The Journal of endocrinology*, *205*(1), 61–68. https://doi.org/10.1677/JOE-09-0337

Faggiano, F., Partanen, T., Kogevinas, M., & Boffetta, P. (1997). Socioeconomic differences in cancer incidence and mortality. *IARC scientific publications*, (138), 65–176.

Fan, S., Meng, Q., Gao, B., Grossman, J., Yadegari, M., Goldberg, I. D., & Rosen, E. M. (2000). Alcohol stimulates estrogen receptor signaling in human breast cancer cell lines. *Cancer research*, *60*(20), 5635–5639.

Fay, M. P., & Freedman, L. S. (1997). Meta-analyses of dietary fats and mammary neoplasms in rodent experiments. *Breast cancer research and treatment*, *46*(2-3), 215–223. https://doi.org/10.1023/a:1005927503865

Feigelson, H. S., & Henderson, B. E. (1996). Estrogens and breast cancer. *Carcinogenesis*, *17*(11), 2279–2284. https://doi.org/10.1093/carcin/17.11.2279

Felmeden, D. C., & Lip, G. Y. (2001). Antihypertensive therapy and cancer risk. *Drug safety*, *24*(10), 727–739. https://doi.org/10.2165/00002018-200124100-00002

Ferrini, K., Ghelfi, F., Mannucci, R., & Titta, L. (2015). Lifestyle, nutrition and breast cancer: facts and presumptions for consideration. *Ecancermedicalscience*, *9*, 557. https://doi.org/10.3332/ecancer.2015.557

Franceschi, S., Favero, A., La Vecchia, C., Barón, A. E., Negri, E., Dal Maso, L., Giacosa, A., Montella, M., Conti, E., & Amadori, D. (1996). Body size indices and breast cancer risk before and after menopause. *International journal of cancer*, *67*(2), 181–186. https://doi.org/10.1002/(SICI)1097-0215(19960717)67:2<181::AID-IJC5>3.0.CO;2-P

Friedenreich, C. M., & Orenstein, M. R. (2002). Physical activity and cancer prevention: etiologic evidence and biological mechanisms. *The Journal of nutrition*, *132*(11 Suppl), 3456S–3464S. https://doi.org/10.1093/jn/132.11.3456S

Frisch, R. E., Gotz-Welbergen, A. V., McArthur, J. W., Albright, T., Witschi, J., Bullen, B., Birnholz, J., Reed, R. B., & Hermann, H. (1981). Delayed menarche and amenorrhea of college athletes in relation to age of onset of training. *JAMA*, *246*(14), 1559–1563.

Furness, D. L., Dekker, G. A., & Roberts, C. T. (2011). DNA damage and health in pregnancy. *Journal of reproductive immunology*, *89*(2), 153–162. https://doi.org/10.1016/j.jri.2011.02.004

Gandini, S., Merzenich, H., Robertson, C., & Boyle, P. (2000). Meta-analysis of studies on breast cancer risk and diet: the role of fruit and vegetable consumption and the intake of associated micronutrients. *European journal of cancer (Oxford, England : 1990)*, *36*(5), 636–646. https://doi.org/10.1016/s0959-8049(00)00022-8

Gapstur, S. M., Potter, J. D., Sellers, T. A., & Folsom, A. R. (1992). Increased risk of breast cancer with alcohol consumption in postmenopausal women. *American journal of epidemiology*, *136*(10), 1221–1231. https://doi.org/10.1093/oxfordjournals.aje.a116430

García-Closas, M., Kelsey, K. T., Hankinson, S. E., Spiegelman, D., Springer, K., Willett, W. C., Speizer, F. E., & Hunter, D. J. (1999). Glutathione S-transferase mu and theta polymorphisms and breast cancer susceptibility. *Journal of the National Cancer Institute*, *91*(22), 1960–1964. https://doi.org/10.1093/jnci/91.22.1960

Garland, M., Morris, J. S., Colditz, G. A., Stampfer, M. J., Spate, V. L., Baskett, C. K., Rosner, B., Speizer, F. E., Willett, W. C., & Hunter, D. J. (1996). Toenail trace element levels and breast cancer: a prospective study. *American journal of epidemiology*, *144*(7), 653–660. https://doi.org/10.1093/oxfordjournals.aje.a008977

Garner, C. E., Burka, L. T., Etheridge, A. E., & Matthews, H. B. (2000). Catechol metabolites of polychlorinated biphenyls inhibit the catechol-O-methyltransferase-mediated metabolism of catechol estrogens. *Toxicology and applied pharmacology*, *162*(2), 115–123. https://doi.org/10.1006/taap.1999.8823

Garner, C. E., Matthews, H. B., & Burka, L. T. (2000). Phenolphthalein metabolite inhibits catechol-O-methyltransferase-mediated metabolism of catechol estrogens: a possible mechanism for carcinogenicity. *Toxicology and applied pharmacology*, *162*(2), 124–131. https://doi.org/10.1006/taap.1999.8830

Garte, S., Gaspari, L., Alexandrie, A. K., Ambrosone, C., Autrup, H., Autrup, J. L., Baranova, H., Bathum, L., Benhamou, S., Boffetta, P., Bouchardy, C.,

Breskvar, K., Brockmoller, J., Cascorbi, I., Clapper, M. L., Coutelle, C., Daly, A., Dell'Omo, M., Dolzan, V., Dresler, C. M., … Taioli, E. (2001). Metabolic gene polymorphism frequencies in control populations. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *10*(12), 1239–1248.

Gasco, M., Yulug, I. G., & Crook, T. (2003). TP53 mutations in familial breast cancer: functional aspects. *Human mutation*, *21*(3), 301–306. https://doi.org/10.1002/humu.10173

Gaudet, M. M., Gapstur, S. M., Sun, J., Diver, W. R., Hannan, L. M., & Thun, M. J. (2013). Active smoking and breast cancer risk: original cohort data and meta-analysis. *Journal of the National Cancer Institute*, *105*(8), 515–525. https://doi.org/10.1093/jnci/djt023

Ghatak, S., Lallawmzuali, D., Lalmawia, Sapkota, R., Zothanpuia, Pautu, J. L., Muthukumaran, R. B., & Senthil Kumar, N. (2014). Mitochondrial D-loop and cytochrome oxidase C subunit I polymorphisms among the breast cancer patients of Mizoram, Northeast India. *Current genetics*, *60*(3), 201–212. https://doi.org/10.1007/s00294-014-0425-2

Gilliam, A. G. (1951). Fertility and cancer of the breast and of the uterine cervix; comparisons between rates of pregnancy in women with cancer at these and other sites. *Journal of the National Cancer Institute*, *12*(2), 287–304.

Glez-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., & Posada, D. (2010). ALTER: program-oriented conversion of DNA and protein alignments. Nucleic acids research, 38(Web Server issue), W14–W18. https://doi.org/10.1093/nar/gkq321

GLOBOCAN: Global Cancer Statistics. Available at: https://www.uicc.org/news/globocan-2020-new-global-cancer-data

Goldberg, M. S., & Labrèche, F. (1996). Occupational risk factors for female breast cancer: a review. *Occupational and environmental medicine*, *53*(3), 145–156. https://doi.org/10.1136/oem.53.3.145

Goss, P. E., & Sierra, S. (1998). Current perspectives on radiation-induced breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, *16*(1), 338–347. https://doi.org/10.1200/JCO.1998.16.1.338

Grodstein, F., Stampfer, M. J., Colditz, G. A., Willett, W. C., Manson, J. E., Joffe, M., Rosner, B., Fuchs, C., Hankinson, S. E., Hunter, D. J., Hennekens, C. H., & Speizer, F. E. (1997). Postmenopausal hormone therapy and mortality. *The New England journal of medicine*, *336*(25), 1769–1775. https://doi.org/10.1056/NEJM199706193362501

Grossman, E., Messerli, F. H., & Goldbourt, U. (2001). Antihypertensive therapy and the risk of malignancies. *European heart journal*, *22*(15), 1343–1352. https://doi.org/10.1053/euhj.2001.2729

Grossman, L. I., & Shoubridge, E. A. (1996). Mitochondrial genetics and human disease. *BioEssays : news and reviews in molecular, cellular and developmental biology*, *18*(12), 983–991. https://doi.org/10.1002/bies.950181208

Ha, M., Mabuchi, K., Sigurdson, A. J., Freedman, D. M., Linet, M. S., Doody, M. M., & Hauptmann, M. (2007). Smoking cigarettes before first childbirth and risk of breast cancer. *American journal of epidemiology*, *166*(1), 55–61. https://doi.org/10.1093/aje/kwm045

Hall, I. J., Newman, B., Millikan, R. C., & Moorman, P. G. (2000). Body size and breast cancer risk in black women and white women: the Carolina Breast Cancer Study. *American journal of epidemiology*, *151*(8), 754–764. https://doi.org/10.1093/oxfordjournals.aje.a010275

Hall, J. M., Lee, M. K., Newman, B., Morrow, J. E., Anderson, L. A., Huey, B., & King, M. C. (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. *Science (New York, N.Y.)*, *250*(4988), 1684–1689. https://doi.org/10.1126/science.2270482

Hallier, E., Langhof, T., Dannappel, D., Leutbecher, M., Schröder, K., Goergens, H. W., Müller, A., & Bolt, H. M. (1993). Polymorphism of glutathione conjugation of methyl bromide, ethylene oxide and dichloromethane in human blood: influence on the induction of sister chromatid exchanges (SCE) in lymphocytes. *Archives of toxicology*, *67*(3), 173–178. https://doi.org/10.1007/BF01973304

Hamajima, N., Hirose, K., Tajima, K., Rohan, T., Calle, E. E., Heath, C. W., Jr, Coates, R. J., Liff, J. M., Talamini, R., Chantarakul, N., Koetsawang, S., Rachawat, D., Morabia, A., Schuman, L., Stewart, W., Szklo, M., Bain, C., Schofield, F., Siskind, V., Band, P., … Collaborative Group on Hormonal Factors in Breast Cancer (2002). Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *British journal of cancer*, *87*(11), 1234–1245. https://doi.org/10.1038/sj.bjc.6600596

Han, H., Guo, W., Shi, W., Yu, Y., Zhang, Y., Ye, X., & He, J. (2017). Hypertension and breast cancer risk: a systematic review and meta-analysis. *Scientific reports*, *7*, 44877. https://doi.org/10.1038/srep44877

Hardefeldt, P. J., Edirimanne, S., & Eslick, G. D. (2012). Diabetes increases the risk of breast cancer: a meta-analysis. *Endocrine-related cancer*, *19*(6), 793–803. https://doi.org/10.1530/ERC-12-0242

Hayes, J. D., Flanagan, J. U., & Jowsey, I. R. (2005). Glutathione transferases. *Annual review of pharmacology and toxicology*, *45*, 51–88. https://doi.org/10.1146/annurev.pharmtox.45.120403.095857

Hein, D. W. (2000). N-Acetyltransferase genetics and their role in predisposition to aromatic and heterocyclic amine-induced carcinogenesis. *Toxicology letters*, *112-113*, 349–356. https://doi.org/10.1016/s0378-4274(99)00226-x

Hein, D. W., Doll, M. A., Fretland, A. J., Leff, M. A., Webb, S. J., Xiao, G. H., Devanaboyina, U. S., Nangju, N. A., & Feng, Y. (2000). Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *9*(1), 29–42.

Helzlsouer, K. J., Selmin, O., Huang, H. Y., Strickland, P. T., Hoffman, S., Alberg, A. J., Watson, M., Comstock, G. W., & Bell, D. (1998). Association between glutathione S-transferase M1, P1, and T1 genetic polymorphisms and development of breast cancer. *Journal of the National Cancer Institute*, *90*(7), 512–518. https://doi.org/10.1093/jnci/90.7.512

Hemminki, A., Markie, D., Tomlinson, I., Avizienyte, E., Roth, S., Loukola, A., Bignell, G., Warren, W., Aminoff, M., Höglund, P., Järvinen, H., Kristo, P., Pelin, K., Ridanpää, M., Salovaara, R., Toro, T., Bodmer, W., Olschwang, S., Olsen, A. S., Stratton, M. R., … Aaltonen, L. A. (1998). A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature*, *391*(6663), 184–187. https://doi.org/10.1038/34432

Henderson, B. E., Ross, R. K., Judd, H. L., Krailo, M. D., & Pike, M. C. (1985). Do regular ovulatory cycles increase breast cancer risk?. *Cancer*, *56*(5), 1206–1208. https://doi.org/10.1002/1097-0142(19850901)56:5<1206::aid-cncr2820560541>3.0.co;2-9

HMTVAR: A Human Mitochondrial Variability and Pathogenicity Database. Available at: https://www.hmtvar.uniba.it/

Holli, K., Isola, J., & Cuzick, J. (1997). Hormone replacement therapy and biological aggressiveness of breast cancer. *Lancet (London, England)*, *350*(9092), 1704–1705. https://doi.org/10.1016/S0140-6736(05)64308-5

Holmberg, L., Baron, J. A., Byers, T., Wolk, A., Ohlander, E. M., Zack, M., & Adami, H. O. (1995). Alcohol intake and breast cancer risk: effect of exposure from 15 years of age. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *4*(8), 843–847.

Holmes, M. D., Liu, S., Hankinson, S. E., Colditz, G. A., Hunter, D. J., & Willett, W. C. (2004). Dietary carbohydrates, fiber, and breast cancer risk. *American journal of epidemiology*, *159*(8), 732–739. https://doi.org/10.1093/aje/kwh112

Horn-Ross, P. L., Canchola, A. J., West, D. W., Stewart, S. L., Bernstein, L., Deapen, D., Pinder, R., Ross, R. K., Anton-Culver, H., Peel, D., Ziogas, A., Reynolds, P., & Wright, W. (2004). Patterns of alcohol consumption and breast cancer risk in the California Teachers Study cohort. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *13*(3), 405–411.

Horwich, A., & Swerdlow, A. J. (2004). Second primary breast cancer after Hodgkin's disease. *British journal of cancer*, *90*(2), 294–298. https://doi.org/10.1038/sj.bjc.6601499

Hsieh, C. C., Trichopoulos, D., Katsouyanni, K., & Yuasa, S. (1990). Age at menarche, age at menopause, height and obesity as risk factors for breast cancer: associations and interactions in an international case-control study. *International journal of cancer*, *46*(5), 796–800. https://doi.org/10.1002/ijc.2910460508

Huang, Z., Hankinson, S. E., Colditz, G. A., Stampfer, M. J., Hunter, D. J., Manson, J. E., Hennekens, C. H., Rosner, B., Speizer, F. E., & Willett, W. C. (1997). Dual effects of weight and weight gain on breast cancer risk. *JAMA*, *278*(17), 1407–1411.

Hunter, D. J., & Willett, W. C. (1993). Diet, body size, and breast cancer. *Epidemiologic reviews*, *15*(1), 110–132. https://doi.org/10.1093/oxfordjournals.epirev.a036096

Hunter, D. J., Spiegelman, D., Adami, H. O., Beeson, L., van den Brandt, P. A., Folsom, A. R., Fraser, G. E., Goldbohm, R. A., Graham, S., & Howe, G. R. (1996). Cohort studies of fat intake and the risk of breast cancer--a pooled analysis. *The New England journal of medicine*, *334*(6), 356–361. https://doi.org/10.1056/NEJM199602083340603

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (2004). Betel-quid and areca-nut chewing and some areca-nut derived nitrosamines. *IARC monographs on the evaluation of carcinogenic risks to humans*, *85*, 1–334.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (1997a). *IARC monographs on the evaluation of carcinogenic risks to humans*. International Agency for Research on Cancer, Lyon, No. 68.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (1997b). Polychlorinated Dibenzo-para-dioxins and Polychlorinated Dibenzofurans. *IARC monographs on the evaluation of carcinogenic risks to humans*. International Agency for Research on Cancer, Lyon, No. 69.

IARCWorking Group on the Evaluation of Carcinogenic Risks to Humans. (1987). Overall evaluations of carcinogenicity: an updating of IARC monographs. *IARC monographs on the evaluation of carcinogenic risks to humans.* International Agency for Research on Cancer, Lyon, volumes 1 to 42.

IARCWorking Group on the Evaluation of Carcinogenic Risks to Humans. (1993). Beryllium, cadmium, mercury, and exposures in the glass manufacturing industry. *IARC monographs on the evaluation of carcinogenic risks to humans*, 58, 1–415.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (2012). Chemical agents and related occupations. *IARC monographs on the evaluation of carcinogenic risks to humans*, 100(Pt F):9-562. PMID: 23189753; PMCID: PMC4781612.

Ibrahim, E. M., & Al-Homaidh, A. (2011). Physical activity and survival after breast cancer diagnosis: meta-analysis of published studies. *Medical oncology (Northwood, London, England)*, *28*(3), 753–765. https://doi.org/10.1007/s12032-010-9536-x

Imagawa, W., Bandyopadhyay, G. K., & Nandi, S. (1990). Regulation of mammary epithelial cell growth in mice and rats. *Endocrine reviews*, *11*(4), 494–523. https://doi.org/10.1210/edrv-11-4-494

Ishii, T., Matsuse, T., Teramoto, S., Matsui, H., Miyao, M., Hosoi, T., Takahashi, H., Fukuchi, Y., & Ouchi, Y. (1999). Glutathione S-transferase P1 (GSTP1)

polymorphism in patients with chronic obstructive pulmonary disease. *Thorax*, *54*(8), 693–696. https://doi.org/10.1136/thx.54.8.693

Islam, D., Islam, M. S., & Jesmin (2022). Association of hypertension, hyperlipidemia, obesity, and demographic risk factors with breast cancer in Bangladeshi women. *Medicine*, 101(46), e31698. https://doi.org/10.1097/MD.0000000000031698

Islami, F., Goding Sauer, A., Miller, K. D., Siegel, R. L., Fedewa, S. A., Jacobs, E. J., McCullough, M. L., Patel, A. V., Ma, J., Soerjomataram, I., Flanders, W. D., Brawley, O. W., Gapstur, S. M., & Jemal, A. (2018). Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA: a cancer journal for clinicians*, *68*(1), 31–54. https://doi.org/10.3322/caac.21440

Janerich, D. T., & Hoff, M. B. (1982). Evidence for a crossover in breast cancer risk factors. *American journal of epidemiology*, *116*(5), 737–742. https://doi.org/10.1093/oxfordjournals.aje.a113462

Johansson, A.S., & Mannervik, B. (2001). Interindividual variability of glutathione transferase expression in man. In Pacifici, G.M. and Pelkonen, O. (eds) *Interindividual Variability in Drug Metabolism in Man*. Taylor & Francis, London

John, E. M., & Kelsey, J. L. (1993). Radiation and other environmental exposures and breast cancer. *Epidemiologic reviews*, *15*(1), 157–162. https://doi.org/10.1093/oxfordjournals.epirev.a036099

Johns, D. R., (1995). Seminars in medicine of the Beth Israel Hospital, Boston. Mitochondrial DNA and disease. *The New England journal of medicine*, *333*(10), 638–644. https://doi.org/10.1056/NEJM199509073331007

Jussawalla, D. J., Yeole, B. B., & Natekar, M. V. (1981). Histological and epidemiological features of breast cancer in different religious groups in greater Bombay. *Journal of surgical oncology*, *18*(3), 269–279. https://doi.org/10.1002/jso.2930180309

Kabuto, M., Akiba, S., Stevens, R. G., Neriishi, K., & Land, C. E. (2000). A prospective study of estradiol and breast cancer in Japanese women. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *9*(6), 575–579.

Kadenbach, B., Arnold, S., Lee, I., & Hüttemann, M. (2004). The possible role of cytochrome c oxidase in stress-induced apoptosis and degenerative diseases. *Biochimica et biophysica acta*, *1655*(1-3), 400–408. https://doi.org/10.1016/j.bbabio.2003.06.005

Kaushal, M., Mishra, A. K., Raju, B. S., Ihsan, R., Chakraborty, A., Sharma, J., Zomawia, E., Verma, Y., Kataki, A., Kapur, S., & Saxena, S. (2010). Betel quid chewing as an environmental risk factor for breast cancer. *Mutation research*, *703*(2), 143–148. https://doi.org/10.1016/j.mrgentox.2010.08.011

Kaushal, M., Mishra, A. K., Sharma, J., Zomawia, E., Kataki, A., Kapur, S., & Saxena, S. (2012). Genomic alterations in breast cancer patients in betel quid and non-betel quid chewers. *PloS one*, *7*(8), e43789. https://doi.org/10.1371/journal.pone.0043789

Kelsey, J. L., Gammon, M. D., & John, E. M. (1993). Reproductive factors and breast cancer. *Epidemiologic reviews*, *15*(1), 36–47. https://doi.org/10.1093/oxfordjournals.epirev.a036115

Kelsey, K. T., Hankinson, S. E., Colditz, G. A., Springer, K., Garcia-Closas, M., Spiegelman, D., Manson, J. E., Garland, M., Stampfer, M. J., Willett, W. C., Speizer, F. E., & Hunter, D. J. (1997). Glutathione S-transferase

class mu deletion polymorphism and breast cancer: results from prevalent versus incident cases. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *6*(7), 511–515.

Kendall, A., Folkerd, E. J., & Dowsett, M. (2007). Influences on circulating oestrogens in postmenopausal women: relationship with breast cancer. *The Journal of steroid biochemistry and molecular biology*, *103*(2), 99–109. https://doi.org/10.1016/j.jsbmb.2006.07.011

Key, T. J. (1999). Serum oestradiol and breast cancer risk. *Endocrine-related cancer*, *6*(2), 175–180. https://doi.org/10.1677/erc.0.0060175

Key, T. J., Appleby, P. N., Reeves, G. K., Roddam, A., Dorgan, J. F., Longcope, C., Stanczyk, F. Z., Stephenson, H. E., Jr, Falk, R. T., Miller, R., Schatzkin, A., Allen, D. S., Fentiman, I. S., Key, T. J., Wang, D. Y., Dowsett, M., Thomas, H. V., Hankinson, S. E., Toniolo, P., Akhmedkhanov, A., … Endogenous Hormones Breast Cancer Collaborative Group (2003). Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *Journal of the National Cancer Institute*, *95*(16), 1218–1226. https://doi.org/10.1093/jnci/djg022

Key, T. J., Verkasalo, P. K., & Banks, E. (2001). Epidemiology of breast cancer. *The Lancet. Oncology*, *2*(3), 133–140. https://doi.org/10.1016/S1470-2045(00)00254-0

Khalis, M., Charbotel, B., Chajès, V., Rinaldi, S., Moskal, A., Biessy, C., Dossus, L., Huybrechts, I., Fort, E., Mellas, N., Elfakir, S., Charaka, H., Nejjari, C., Romieu, I., & El Rhazi, K. (2018). Menstrual and reproductive factors and risk of breast cancer: A case-control study in the Fez region, Morocco. *PloS one*, *13*(1), e0191333. https://doi.org/10.1371/journal.pone.0191333

Kheifets, L. I., & Matkin, C. C. (1999). Industrialization, electromagnetic fields, and breast cancer risk. *Environmental health perspectives*, *107 Suppl 1*(Suppl 1), 145–154. https://doi.org/10.1289/ehp.99107s1145

Khuder, S. A., Mutgi, A. B., & Nugent, S. (2001). Smoking and breast cancer: a meta-analysis. *Reviews on environmental health*, *16*(4), 253–261. https://doi.org/10.1515/reveh.2001.16.4.253

Kimi, L., Ghatak, S., Yadav, R. P., Chhuani, L., Lallawmzuali, D., Pautu, J. L., & Kumar, N. S. (2016). Relevance of GSTM1, GSTT1 and GSTP1 Gene Polymorphism to Breast Cancer Susceptibility in Mizoram Population, Northeast India. *Biochemical genetics*, *54*(1), 41–49. https://doi.org/10.1007/s10528-015-9698-5

Ko, L. J., & Prives, C. (1996). p53: puzzle and paradigm. *Genes & development*, *10*(9), 1054–1072. https://doi.org/10.1101/gad.10.9.1054

Kotepui, M. (2016). Diet and risk of breast cancer. *Contemporary Oncology/ Współczesna Onkologia*, *20*(1), 13-19.

Kuper, H., Ye, W., Weiderpass, E., Ekbom, A., Trichopoulos, D., Nyrén, O., & Adami, H. O. (2000). Alcohol and breast cancer risk: the alcoholism paradox. *British journal of cancer*, *83*(7), 949–951. https://doi.org/10.1054/bjoc.2000.1360

Kvåle, G., & Heuch, I. (1988). Menstrual factors and breast cancer risk. *Cancer*, *62*(8), 1625–1631. https://doi.org/10.1002/1097-0142(19881015)62:8<1625::aid-cncr2820620828>3.0.co;2-k

Laidlaw, I. J., Clarke, R. B., Howell, A., Owen, A. W., Potten, C. S., & Anderson, E. (1995). The proliferation of normal human breast tissue implanted into athymic nude mice is stimulated by estrogen but not progesterone. *Endocrinology*, *136*(1), 164–171. https://doi.org/10.1210/endo.136.1.7828527

Land, C. E., Tokunaga, M., Koyama, K., Soda, M., Preston, D. L., Nishimori, I., & Tokuoka, S. (2003). Incidence of female breast cancer among atomic bomb survivors, Hiroshima and Nagasaki, 1950-1990. *Radiation research*, *160*(6), 707–717. https://doi.org/10.1667/rr3082

Landi, S. (2000). Mammalian class theta GST and differential susceptibility to carcinogens: a review. *Mutation research*, *463*(3), 247–283. https://doi.org/10.1016/s1383-5742(00)00050-8

Lane-Claypon, J.E. (1926). A further report on cancer of the breast, with special reference to its associated antecedent conditions. *Ministry of Health Reports on public health, no 23,* London.

Largent, J. A., McEligot, A. J., Ziogas, A., Reid, C., Hess, J., Leighton, N., Peel, D., & Anton-Culver, H. (2006). Hypertension, diuretics and breast cancer risk. *Journal of human hypertension*, *20*(10), 727–732. https://doi.org/10.1038/sj.jhh.1002075

Layde, P. M., Webster, L. A., Baughman, A. L., Wingo, P. A., Rubin, G. L., & Ory, H. W. (1989). The independent associations of parity, age at first full term pregnancy, and duration of breastfeeding with the risk of breast cancer. Cancer and Steroid Hormone Study Group. *Journal of clinical epidemiology*, *42*(10), 963–973. https://doi.org/10.1016/0895-4356(89)90161-3

Lee, I. M. (1999). Antioxidant vitamins in the prevention of cancer. *Proceedings of the Association of American Physicians*, *111*(1), 10–15. https://doi.org/10.1046/j.1525-1381.1999.09230.x

Lee, I., Bender, E., Arnold, S., & Kadenbach, B. (2001). New control of mitochondrial membrane potential and ROS formation--a hypothesis. *Biological chemistry*, *382*(12), 1629–1636. https://doi.org/10.1515/BC.2001.198

Lewis-Michl, E. L., Melius, J. M., Kallenbach, L. R., Ju, C. L., Talbot, T. O., Orr, M. F., & Lauridsen, P. E. (1996). Breast cancer risk and residence near industry or traffic in Nassau and Suffolk Counties, Long Island, New York. *Archives of environmental health*, *51*(4), 255–265. https://doi.org/10.1080/00039896.1996.9936024

Li, C. I., Chlebowski, R. T., Freiberg, M., Johnson, K. C., Kuller, L., Lane, D., Lessin, L., O'Sullivan, M. J., Wactawski-Wende, J., Yasmeen, S., & Prentice, R. (2010). Alcohol consumption and risk of postmenopausal breast cancer by subtype: the women's health initiative observational study. *Journal of the National Cancer Institute*, *102*(18), 1422–1431. https://doi.org/10.1093/jnci/djq316

Li, C. I., Malone, K. E., Daling, J. R., Potter, J. D., Bernstein, L., Marchbanks, P. A., Strom, B. L., Simon, M. S., Press, M. F., Ursin, G., Burkman, R. T., Folger, S. G., Norman, S., McDonald, J. A., & Spirtas, R. (2008). Timing of menarche and first full-term birth in relation to breast cancer risk. *American journal of epidemiology*, *167*(2), 230–239. https://doi.org/10.1093/aje/kwm271

Li, D., Wang, M., Dhingra, K., & Hittelman, W. N. (1996). Aromatic DNA adducts in adjacent tissues of breast cancer patients: clues to breast cancer etiology. *Cancer research*, *56*(2), 287–293.

Li, L., Zhou, X., & Wang, Y. (2018). Smokeless tobacco extract inhibits proliferation and promotes apoptosis in oral mucous fibroblasts. *Oncology letters*, *16*(4), 5066–5074. https://doi.org/10.3892/ol.2018.9252

Li, X., You, R., Wang, X., Liu, C., Xu, Z., Zhou, J., Yu, B., Xu, T., Cai, H., & Zou, Q. (2016). Effectiveness of Prophylactic Surgeries in BRCA1 or BRCA2 Mutation Carriers: A Meta-analysis and Systematic Review. *Clinical*

cancer research: an official journal of the American Association for Cancer Research, 22(15), 3971–3981. https://doi.org/10.1158/1078-0432.CCR-15-1465

Lindgren, A., Pukkala, E., Tuomilehto, J., & Nissinen, A. (2007). Incidence of breast cancer among postmenopausal, hypertensive women. *International journal of cancer*, *121*(3), 641–644. https://doi.org/10.1002/ijc.22689

Lindholm, L. H., Anderson, H., Ekbom, T., Hansson, L., Lanke, J., Dahlöf, B., de Faire, U., Forsén, K., Hedner, T., Linjer, E., Scherstén, B., Wester, P., & Möller, T. (2001). Relation between drug treatment and cancer in hypertensives in the Swedish Trial in Old Patients with Hypertension 2: a 5-year, prospective, randomised, controlled trial. *Lancet (London, England)*, *358*(9281), 539–544. https://doi.org/10.1016/s0140-6736(01)05704-x

Lipworth, L., Bailey, L. R., & Trichopoulos, D. (2000). History of breast-feeding in relation to breast cancer risk: a review of the epidemiologic literature. *Journal of the National Cancer Institute*, *92*(4), 302–312. https://doi.org/10.1093/jnci/92.4.302

Lizard-Nacol, S., Coudert, B., Colosetti, P., Riedinger, J. M., Fargeot, P., & Brunet-Lecomte, P. (1999). Glutathione S-transferase M1 null genotype: lack of association with tumour characteristics and survival in advanced breast cancer. *Breast cancer research: BCR*, *1*(1), 81–87. https://doi.org/10.1186/bcr17

London, S. J., Pogoda, J. M., Hwang, K. L., Langholz, B., Monroe, K. R., Kolonel, L. N., Kaune, W. T., Peters, J. M., & Henderson, B. E. (2003). Residential magnetic field exposure and breast cancer risk: a nested case-control study from a multiethnic cohort in Los Angeles County, California. *American journal of epidemiology*, *158*(10), 969–980. https://doi.org/10.1093/aje/kwg254

Lorincz, A. M., & Sukumar, S. (2006). Molecular links between obesity and breast cancer. *Endocrine-related cancer*, *13*(2), 279–292. https://doi.org/10.1677/erc.1.00729

Lubin, F., Ruder, A. M., Wax, Y., & Modan, B. (1985). Overweight and changes in weight throughout adult life in breast cancer etiology. A case-control study. *American journal of epidemiology*, *122*(4), 579–588. https://doi.org/10.1093/oxfordjournals.aje.a114137

Lux, M. P., Fasching, P. A., & Beckmann, M. W. (2006). Hereditary breast and ovarian cancer: review and future perspectives. *Journal of molecular medicine (Berlin, Germany)*, *84*(1), 16–28. https://doi.org/10.1007/s00109-005-0696-7

Lynch, H. T., Silva, E., Snyder, C., & Lynch, J. F. (2008). Hereditary breast cancer: part I. Diagnosing hereditary breast cancer syndromes. *The breast journal*, *14*(1), 3–13. https://doi.org/10.1111/j.1524-4741.2007.00515.x

MacMahon, B., Cole, P., & Brown, J. (1973). Etiology of human breast cancer: a review. *Journal of the National Cancer Institute*, *50*(1), 21–42. https://doi.org/10.1093/jnci/50.1.21

MacMahon, B., Cole, P., Lin, T. M., Lowe, C. R., Mirra, A. P., Ravnihar, B., Salber, E. J., Valaoras, V. G., & Yuasa, S. (1970). Age at first birth and breast cancer risk. *Bulletin of the World Health Organization*, *43*(2), 209–221.

Mahanta, J., Chetia, M., Hazarika, N.C., Narain, K., & Sharma, S.K. (1998). Toxicity of tuibur, a unique form of tobacco smoke extract used in Mizoram, India. *Current Science*, 75(4), 381-384.

Mahdavi, M., Nassiri, M., Kooshyar, M. M., Vakili-Azghandi, M., Avan, A., Sandry, R., Pillai, S., Lam, A. K., & Gopalan, V. (2019). Hereditary breast cancer;

Genetic penetrance and current status with BRCA. *Journal of cellular physiology*, *234*(5), 5741–5750. https://doi.org/10.1002/jcp.27464

Malakar, M., Devi, K. R., Phukan, R. K., Kaur, T., Deka, M., Puia, L., Barua, D., Mahanta, J., & Narain, K. (2012). Genetic polymorphism of glutathione S-transferases M1 and T1, tobacco habits and risk of stomach cancer in Mizoram, India. Asian Pacific journal of cancer prevention : APJCP, 13(9), 4725–4732. https://doi.org/10.7314/apjcp.2012.13.9.4725

Malik, F. A., Sanders, A. J., Jones, A. D., Mansel, R. E., & Jiang, W. G. (2009). Transcriptional and translational modulation of KAI1 expression in ductal carcinoma of the breast and the prognostic significance. *International journal of molecular medicine*, *23*(2), 273–278.\

Manjer, J., Kaaks, R., Riboli, E., & Berglund, G. (2001). Risk of breast cancer in relation to anthropometry, blood pressure, blood lipids and glucose metabolism: a prospective study within the Malmö Preventive Project. *European journal of cancer prevention: the official journal of the European Cancer Prevention Organisation (ECP)*, *10*(1), 33–42. https://doi.org/10.1097/00008469-200102000-00004

Mannervik, B., Board, P. G., Hayes, J. D., Listowsky, I., & Pearson, W. R. (2005). Nomenclature for mammalian soluble glutathione transferases. *Methods in enzymology*, *401*, 1–8. https://doi.org/10.1016/S0076-6879(05)01001-3

Männistö, S., Pietinen, P., Pyy, M., Palmgren, J., Eskelinen, M., & Uusitupa, M. (1996). Body-size indicators and risk of breast cancer according to menopause and estrogen-receptor status. *International journal of cancer*, *68*(1), 8–13. https://doi.org/10.1002/(SICI)1097-0215(19960927)68:1<8::AID-IJC2>3.0.CO;2-V

Marcus, P. M., Newman, B., Millikan, R. C., Moorman, P. G., Baird, D. D., & Qaqish, B. (2000). The associations of adolescent cigarette smoking, alcoholic beverage consumption, environmental tobacco smoke, and ionizing radiation with subsequent breast cancer risk (United States). *Cancer causes & control: CCC*, *11*(3), 271–278. https://doi.org/10.1023/a:1008911902994

Mathur, P., Sathishkumar, K., Chaturvedi, M., Das, P., Sudarshan, K. L., Santhappan, S., Nallasamy, V., John, A., Narasimhan, S., Roselind, F. S., & ICMR-NCDIR-NCRP Investigator Group (2020). Cancer Statistics, 2020: Report From National Cancer Registry Programme, India. *JCO global oncology*, *6*, 1063–1075. https://doi.org/10.1200/GO.20.00122

Maurya, A. P., & Brahmachari, S. (2022). Association of hormonal and reproductive risk factors with breast cancer in Indian women: A systematic review of case-control studies. *Indian journal of cancer*, 10.4103/ijc.IJC_271_21. Advance online publication. https://doi.org/10.4103/ijc.IJC_271_21

McCormick, B. (1999). Radiation and Local Control in Early Invasive Breast Cancer. *The breast journal*, *5*(5), 330–334. https://doi.org/10.1046/j.1524-4741.1999.00007.x

McCullough, M. L., Patel, A. V., Kushi, L. H., Patel, R., Willett, W. C., Doyle, C., Thun, M. J., & Gapstur, S. M. (2011). Following cancer prevention guidelines reduces risk of cancer, cardiovascular disease, and all-cause mortality. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *20*(6), 1089–1097. https://doi.org/10.1158/1055-9965.EPI-10-1173

McDowall, M. E. (1986). Mortality of persons resident in the vicinity of electricity transmission facilities. *British journal of cancer*, *53*(2), 271–279. https://doi.org/10.1038/bjc.1986.45

McElroy, J. A., Shafer, M. M., Trentham-Dietz, A., Hampton, J. M., & Newcomb, P. A. (2006). Cadmium exposure and breast cancer risk. *Journal of the National Cancer Institute*, *98*(12), 869–873. https://doi.org/10.1093/jnci/djj233

McKeown, N. (1999). Antioxidants and breast cancer. *Nutrition reviews*, *57*(10), 321–324. https://doi.org/10.1111/j.1753-4887.1999.tb06906.x

McLellan, R. A., Oscarson, M., Alexandrie, A. K., Seidegård, J., Evans, D. A., Rannug, A., & Ingelman-Sundberg, M. (1997). Characterization of a human glutathione S-transferase mu cluster containing a duplicated GSTM1 gene that causes ultrarapid enzyme activity. *Molecular pharmacology*, *52*(6), 958–965. https://doi.org/10.1124/mol.52.6.958

McManus, M. J., & Welsch, C. W. (1984). The effect of estrogen, progesterone, thyroxine, and human placental lactogen on DNA synthesis of human breast ductal epithelium maintained in athymic nude mice. *Cancer*, *54*(9), 1920–1927. https://doi.org/10.1002/1097-0142(19841101)54:9<1920::aid-cncr2820540924>3.0.co;2-f

McPherson, K., Steel, C. M., & Dixon, J. M. (2000). ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *BMJ (Clinical research ed.)*, *321*(7261), 624–628. https://doi.org/10.1136/bmj.321.7261.624

McTiernan, A. (2003). Behavioral risk factors in breast cancer: can risk be modified?. *The oncologist*, *8*(4), 326–334. https://doi.org/10.1634/theoncologist.8-4-326

Megdal, S. P., Kroenke, C. H., Laden, F., Pukkala, E., & Schernhammer, E. S. (2005). Night work and breast cancer risk: a systematic review and meta-analysis. *European journal of cancer (Oxford, England : 1990)*, *41*(13), 2023–2032. https://doi.org/10.1016/j.ejca.2005.05.010

Michels, K. B., & Willett, W. C. (1996). Does induced or spontaneous abortion affect the risk of breast cancer?. *Epidemiology (Cambridge, Mass.)*, *7*(5), 521–528.

Michels, K. B., Solomon, C. G., Hu, F. B., Rosner, B. A., Hankinson, S. E., Colditz, G. A., Manson, J. E., & Nurses' Health Study (2003). Type 2 diabetes and subsequent incidence of breast cancer in the Nurses' Health Study. *Diabetes care*, *26*(6), 1752–1758. https://doi.org/10.2337/diacare.26.6.1752

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., & Ding, W. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science (New York, N.Y.)*, *266*(5182), 66–71. https://doi.org/10.1126/science.7545954

Millikan, R., Pittman, G., Tse, C. K., Savitz, D. A., Newman, B., & Bell, D. (2000). Glutathione S-transferases M1, T1, and P1 and breast cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *9*(6), 567–573.

MITOMAP: A Human Mitochondrial Genome Database. Available at: http://www.mitomap.org

Modica-Napolitano, J. S., Kulawiec, M., & Singh, K. K. (2007). Mitochondria and human cancer. *Current molecular medicine*, *7*(1), 121-131

Morabia, A. (2002). Smoking (active and passive) and breast cancer: epidemiologic evidence up to June 2001. *Environmental and molecular mutagenesis*, *39*(2-3), 89–95. https://doi.org/10.1002/em.10046

Moral, R., Wang, R., Russo, I. H., Lamartiniere, C. A., Pereira, J., & Russo, J. (2008). Effect of prenatal exposure to the endocrine disruptor bisphenol A on mammary gland morphology and gene expression signature. *The Journal of endocrinology*, *196*(1), 101–112. https://doi.org/10.1677/JOE-07-0056

MTDNA: Human Mitochondrial DNA. Available at: https://www.mdpi.com/2076-3921/8/9/392

Murray, T. J., Maffini, M. V., Ucci, A. A., Sonnenschein, C., & Soto, A. M. (2007). Induction of mammary gland ductal hyperplasias and carcinoma in situ following fetal bisphenol A exposure. *Reproductive toxicology (Elmsford, N.Y.)*, *23*(3), 383–390. https://doi.org/10.1016/j.reprotox.2006.10.002

Nagrani, R. T., Budukh, A., Koyande, S., Panse, N. S., Mhatre, S. S., & Badwe, R. (2014). Rural urban differences in breast cancer in India. *Indian journal of cancer*, *51*(3), 277–281. https://doi.org/10.4103/0019-509X.146793

Nagrani, R. T., Mhatre, S., Boffetta, P., Rajaraman, P., Badwe, R., Gupta, S., Romieu, I., Parmar, V., & Dikshit, R. (2016). Understanding rural-urban differences in risk factors for breast cancer in an Indian population. *Cancer causes & control : CCC*, *27*(2), 199–208. https://doi.org/10.1007/s10552-015-0697-y

Nair, U., Bartsch, H., & Nair, J. (2004). Alert for an epidemic of oral cancer due to use of the betel quid substitutes gutkha and pan masala: a review of agents and causative mechanisms. *Mutagenesis*, *19*(4), 251–262. https://doi.org/10.1093/mutage/geh036

Namslauer, I., & Brzezinski, P. (2009). A mitochondrial DNA mutation linked to colon cancer results in proton leaks in cytochrome c oxidase. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(9), 3402–3407. https://doi.org/10.1073/pnas.0811450106

Navarro Silvera, S. A., & Rohan, T. E. (2007). Trace elements and cancer risk: a review of the epidemiologic evidence. *Cancer causes & control : CCC*, *18*(1), 7–27. https://doi.org/10.1007/s10552-006-0057-z

NCDIR: National Cancer Registry Report. Available at: https://ncdirindia.org/All_Reports/Report_2020/default.aspx

NCI: National Cancer Institute. Available at: http:// www.cancer.gov/cancerinfo/ere-workshop-report

Nedelcheva Kristensen, V., Andersen, T. I., Erikstein, B., Geitvik, G., Skovlund, E., Nesland, J. M., & Børresen-Dale, A. L. (1998). Single tube multiplex polymerase chain reaction genotype analysis of GSTM1, GSTT1 and GSTP1: relation of genotypes to TP53 tumor status and clinicopathological variables in breast cancer patients. *Pharmacogenetics*, *8*(5), 441–447. https://doi.org/10.1097/00008571-199810000-00009

Nelen, M. R., Padberg, G. W., Peeters, E. A., Lin, A. Y., van den Helm, B., Frants, R. R., Coulon, V., Goldstein, A. M., van Reen, M. M., Easton, D. F., Eeles, R. A., Hodgsen, S., Mulvihill, J. J., Murday, V. A., Tucker, M. A., Mariman, E. C., Starink, T. M., Ponder, B. A., Ropers, H. H., Kremer, H., … Eng, C. (1996). Localization of the gene for Cowden disease to chromosome 10q22-23. *Nature genetics*, *13*(1), 114–116. https://doi.org/10.1038/ng0596-114

Nelson, H. H., Wiencke, J. K., Christiani, D. C., Cheng, T. J., Zuo, Z. F., Schwartz, B. S., Lee, B. K., Spitz, M. R., Wang, M., & Xu, X. (1995). Ethnic differences in the prevalence of the homozygous deleted genotype of glutathione S-transferase theta. *Carcinogenesis*, *16*(5), 1243–1245. https://doi.org/10.1093/carcin/16.5.1243

Newcomb, P. A., Klein, R., Klein, B. E., Haffner, S., Mares-Perlman, J., Cruickshanks, K. J., & Marcus, P. M. (1995). Association of dietary and life-style factors

with sex hormones in postmenopausal women. *Epidemiology (Cambridge, Mass.)*, *6*(3), 318–321. https://doi.org/10.1097/00001648-199505000-00022

Newcomb, P. A., Storer, B. E., Longnecker, M. P., Mittendorf, R., Greenberg, E. R., Clapp, R. W., Burke, K. P., Willett, W. C., & MacMahon, B. (1994). Lactation and a reduced risk of premenopausal breast cancer. *The New England journal of medicine*, *330*(2), 81–87. https://doi.org/10.1056/NEJM199401133300201

NHP: National Health Portal. Available at: https://www.nhp.gov.in/healthlyliving/women-s-health

Nie, J., Beyea, J., Bonner, M. R., Han, D., Vena, J. E., Rogerson, P., Vito, D., Muti, P., Trevisan, M., Edge, S. B., & Freudenheim, J. L. (2007). Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. *Cancer causes & control : CCC*, *18*(9), 947–955. https://doi.org/10.1007/s10552-007-9036-2

Overview of Carcinogenesis. Available at: https://sphweb.bumc.bu.edu

Parise, C. A., Bauer, K. R., Brown, M. M., & Caggiano, V. (2009). Breast cancer subtypes as defined by the estrogen receptor (ER), progesterone receptor (PR), and the human epidermal growth factor receptor 2 (HER2) among women with invasive breast cancer in California, 1999-2004. *The breast journal*, *15*(6), 593–602. https://doi.org/10.1111/j.1524-4741.2009.00822.x

Park, S. K., Yoo, K. Y., Lee, S. J., Kim, S. U., Ahn, S. H., Noh, D. Y., Choe, K. J., Strickland, P. T., Hirvonen, A., & Kang, D. (2000). Alcohol consumption, glutathione S-transferase M1 and T1 genetic polymorphisms and breast cancer risk. *Pharmacogenetics*, *10*(4), 301–309. https://doi.org/10.1097/00008571-200006000-00004

Parkin, D. M. (2004). International variation. *Oncogene*, *23*(38), 6329–6340. https://doi.org/10.1038/sj.onc.1207726

Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global cancer statistics, 2002. *CA: a cancer journal for clinicians*, *55*(2), 74–108. https://doi.org/10.3322/canjclin.55.2.74

Parrella, P., Seripa, D., Matera, M. G., Rabitti, C., Rinaldi, M., Mazzarelli, P., Gravina, C., Gallucci, M., Altomare, V., Flammia, G., Casalino, B., Benedetti-Panici, P. L., & Fazio, V. M. (2003). Mutations of the D310 mitochondrial mononucleotide repeat in primary tumors and cytological specimens. *Cancer letters*, *190*(1), 73–77. https://doi.org/10.1016/s0304-3835(02)00578-5

Parrella, P., Xiao, Y., Fliss, M., Sanchez-Cespedes, M., Mazzarelli, P., Rinaldi, M., Nicol, T., Gabrielson, E., Cuomo, C., Cohen, D., Pandit, S., Spencer, M., Rabitti, C., Fazio, V. M., & Sidransky, D. (2001). Detection of mitochondrial DNA mutations in primary breast cancer and fine-needle aspirates. *Cancer research*, *61*(20), 7623–7626.

Pathak, D. R., & Whittemore, A. S. (1992). Combined effects of body size, parity, and menstrual events on breast cancer incidence in seven countries. *American journal of epidemiology*, *135*(2), 153–168. https://doi.org/10.1093/oxfordjournals.aje.a116268

Pathak, D. R., Osuch, J. R., & He, J. (2000). Breast carcinoma etiology: current knowledge and new insights into the effects of reproductive and hormonal risk factors in black and white populations. *Cancer*, *88*(5 Suppl), 1230–1238. https://doi.org/10.1002/(sici)1097-0142(20000301)88:5+<1230::aid-cncr9>3.0.co;2-f

Pathak, D. R., Speizer, F. E., Willett, W. C., Rosner, B., & Lipnick, R. J. (1986). Parity and breast cancer risk: possible effect on age at diagnosis. *International journal of cancer*, *37*(1), 21–25. https://doi.org/10.1002/ijc.2910370105

Payne, C. M., Holubec, H., Bernstein, C., Bernstein, H., Dvorak, K., Green, S. B., Wilson, M., Dall'Agnol, M., Dvorakova, B., Warneke, J., & Garewal, H. (2005). Crypt-restricted loss and decreased protein expression of cytochrome C oxidase subunit I as potential hypothesis-driven biomarkers of colon cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *14*(9), 2066–2075. https://doi.org/10.1158/1055-9965.EPI-05-0180

PBCR: Population Based Cancer Registry. Available at: https://pbcr.ncdirindia.org/pbcr_2.0/

Pearson, W. R., Vorachek, W. R., Xu, S. J., Berger, R., Hart, I., Vannais, D., & Patterson, D. (1993). Identification of class-mu glutathione transferase genes GSTM1-GSTM5 on human chromosome 1p13. *American journal of human genetics*, *53*(1), 220–233.

Peeters, P. H., van Noord, P. A., Hoes, A. W., Fracheboud, J., Gimbrère, C. H., & Grobbee, D. E. (2000). Hypertension and breast cancer risk in a 19-year follow-up study (the DOM cohort). Diagnostic investigation into mammarian cancer. *Journal of hypertension*, *18*(3), 249–254. https://doi.org/10.1097/00004872-200018030-00002

Pemble, S., Schroeder, K. R., Spencer, S. R., Meyer, D. J., Hallier, E., Bolt, H. M., Ketterer, B., & Taylor, J. B. (1994). Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. *The Biochemical journal*, *300 (Pt 1)*(Pt 1), 271–276. https://doi.org/10.1042/bj3000271

Phukan, R. K., Narain, K., Zomawia, E., Hazarika, N. C., & Mahanta, J. (2006). Dietary habits and stomach cancer in Mizoram, India. *Journal of gastroenterology*, *41*(5), 418–424. https://doi.org/10.1007/s00535-006-1761-x

Phukan, R. K., Saikia, B. J., Borah, P. K., Zomawia, E., Sekhon, G. S., & Mahanta, J. (2014). Role of household exposure, dietary habits and glutathione S-Transferases M1, T1 polymorphisms in susceptibility to lung cancer among women in Mizoram India. *Asian Pacific journal of cancer prevention : APJCP*, *15*(7), 3253–3260. https://doi.org/10.7314/apjcp.2014.15.7.3253

Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. (1995). World Health Organization technical report series, 854, 1–452.

Pike, M. C., Krailo, M. D., Henderson, B. E., Casagrande, J. T., & Hoel, D. G. (1983). 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. *Nature*, *303*(5920), 767–770. https://doi.org/10.1038/303767a0

Pike, M. C., Pearce, C. L., & Wu, A. H. (2004). Prevention of cancers of the breast, endometrium and ovary. *Oncogene*, *23*(38), 6379–6391. https://doi.org/10.1038/sj.onc.1207899

Pope, M. H. (2004). Bernardino Ramazzini: the father of occupational medicine. Spine, 29(20), 2335–2338. https://doi.org/10.1097/01.brs.0000142437.70429.a8

Putignani, L., Raffa, S., Pescosolido, R., Aimati, L., Signore, F., Torrisi, M. R., & Grammatico, P. (2008). Alteration of expression levels of the oxidative phosphorylation system (OXPHOS) in breast cancer cell mitochondria. Breast cancer research and treatment, 110(3), 439–452. https://doi.org/10.1007/s10549-007-9738-x

Raftogianis, R., Creveling, C., Weinshilboum, R., & Weisz, J. (2000). Estrogen metabolism by conjugation. Journal of the National Cancer Institute.

Monographs, (27), 113–124. https://doi.org/10.1093/oxfordjournals.jncimonographs.a024234

Rajbongshi, N., Mahanta, L. B., & Nath, D. C. (2015). Evaluation of Female Breast Cancer Risk Among the Betel Quid Chewer: A Bio-Statistical Assessment in Assam, India. Nepal journal of epidemiology, 5(2), 494–498. https://doi.org/10.3126/nje.v5i2.12832

Rebbeck, T. R. (1999). Inherited genetic predisposition in breast cancer. A population-based perspective. Cancer, 86(11 Suppl), 2493–2501. https://doi.org/10.1002/(sici)1097-0142(19991201)86:11+<2493::aid-cncr6>3.0.co;2-z

Reeves, G. K., Pirie, K., Beral, V., Green, J., Spencer, E., Bull, D., & Million Women Study Collaboration (2007). Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. BMJ (Clinical research ed.), 335(7630), 1134. https://doi.org/10.1136/bmj.39367.495995.AE

Reichman, M. E., Judd, J. T., Longcope, C., Schatzkin, A., Clevidence, B. A., Nair, P. P., Campbell, W. S., & Taylor, P. R. (1993). Effects of alcohol consumption on plasma and urinary hormone concentrations in premenopausal women. Journal of the National Cancer Institute, 85(9), 722–727. https://doi.org/10.1093/jnci/85.9.722

Reif, J. S., Burch, J. B., Pitrat, C. A., Keefe, T. J., & Yost, M. J. (1995). Melatonin levels in electric utility workers. Annual Review of Research on Biological Effects of Electric and Magnetic Fields from the Generation, Delivery and Use of Electricity, Palm Springs, CA, USA, 48-49.

Renehan, A. G., Tyson, M., Egger, M., Heller, R. F., & Zwahlen, M. (2008). Body-mass index and incidence of cancer: a systematic review and meta-analysis of

prospective observational studies. Lancet (London, England), 371(9612), 569–578. https://doi.org/10.1016/S0140-6736(08)60269-X

Rich-Edwards, J. W., Goldman, M. B., Willett, W. C., Hunter, D. J., Stampfer, M. J., Colditz, G. A., & Manson, J. E. (1994). Adolescent body mass index and infertility caused by ovulatory disorder. American journal of obstetrics and gynecology, 171(1), 171–177. https://doi.org/10.1016/0002-9378(94)90465-0

Rigoni-Stern, R. (1842). Fatti statistici relativi alia malattie cancerose. Giornali per Servireal Progressi della Patologia e della Terapeutica, 2:507-17.

Riou, G., Favre, M., Jeannel, D., Bourhis, J., Le Doussal, V., & Orth, G. (1990). Association between poor prognosis in early-stage invasive cervical carcinomas and non-detection of HPV DNA. Lancet (London, England), 335(8699), 1171–1174. https://doi.org/10.1016/0140-6736(90)92693-c

Rohan, T. E., Wong, L. J., Wang, T., Haines, J., & Kabat, G. C. (2010). Do alterations in mitochondrial DNA play a role in breast carcinogenesis?. *Journal of oncology, 2010,* 604304. https://doi.org/10.1155/2010/604304

Romieu, I., Hernández-Avila, M., Lazcano, E., Lopez, L., & Romero-Jaime, R. (1996). Breast cancer and lactation history in Mexican women. American journal of epidemiology, 143(6), 543–552. https://doi.org/10.1093/oxfordjournals.aje.a008784

Ross, R. K., Paganini-Hill, A., Wan, P. C., & Pike, M. C. (2000). Effect of hormone replacement therapy on breast cancer risk: estrogen versus estrogen plus progestin. Journal of the National Cancer Institute, 92(4), 328–332. https://doi.org/10.1093/jnci/92.4.328

Rudel, R. A., Attfield, K. R., Schifano, J. N., & Brody, J. G. (2007). Chemicals causing mammary gland tumors in animals signal new directions for epidemiology,

chemicals testing, and risk assessment for breast cancer prevention. Cancer, 109(12 Suppl), 2635–2666. https://doi.org/10.1002/cncr.22653

Rundle, A., Tang, D., Zhou, J., Cho, S., & Perera, F. (2000). The association between glutathione S-transferase M1 genotype and polycyclic aromatic hydrocarbon-DNA adducts in breast tissue. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 9(10), 1079–1085.

Russo, I. H., & Russo, J. (1993). Physiological bases of breast cancer prevention. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP), 2 Suppl 3, 101–111.

Russo, J., Gusterson, B. A., Rogers, A. E., Russo, I. H., Wellings, S. R., & van Zwieten, M. J. (1990). Comparative study of human and rat mammary tumorigenesis. Laboratory investigation; a journal of technical methods and pathology, 62(3), 244–278.

Russo, J., Hu, Y. F., Yang, X., & Russo, I. H. (2000). Developmental, cellular, and molecular basis of human breast cancer. Journal of the National Cancer Institute. Monographs, (27), 17–37. https://doi.org/10.1093/oxfordjournals.jncimonographs.a024241

Russo, J., Tay, L. K., & Russo, I. H. (1982). Differentiation of the mammary gland and susceptibility to carcinogenesis. Breast cancer research and treatment, 2(1), 5–73. https://doi.org/10.1007/BF01805718

Ryberg, D., Skaug, V., Hewer, A., Phillips, D. H., Harries, L. W., Wolf, C. R., Ogreid, D., Ulvik, A., Vu, P., & Haugen, A. (1997). Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct

levels and cancer risk. Carcinogenesis, 18(7), 1285–1289. https://doi.org/10.1093/carcin/18.7.1285

Safe, S. H. (1997). Is there an association between exposure to environmental estrogens and breast cancer?. Environmental health perspectives, 105 Suppl 3(Suppl 3), 675–678. https://doi.org/10.1289/ehp.97105s3675

Santodonato, J. (1997). Review of the estrogenic and antiestrogenic activity of polycyclic aromatic hydrocarbons: relationship to carcinogenicity. Chemosphere, 34(4), 835–848. https://doi.org/10.1016/s0045-6535(97)00012-x

Santos, G. C., Jr, Góes, A. C., Vitto, H. d., Moreira, C. C., Avvad, E., Rumjanek, F. D., & Moura Gallo, C. V. (2012). Genomic instability at the 13q31 locus and somatic mtDNA mutation in the D-loop site correlate with tumor aggressiveness in sporadic Brazilian breast cancer cases. Clinics (Sao Paulo, Brazil), 67(10), 1181–1190. https://doi.org/10.6061/clinics/2012(10)10

Schairer, C., Lubin, J., Troisi, R., Sturgeon, S., Brinton, L., & Hoover, R. (2000). Menopausal estrogen and estrogen-progestin replacement therapy and breast cancer risk. *JAMA*, *283*(4), 485–491. https://doi.org/10.1001/jama.283.4.485

Schatz, G., (1996). The protein import system of mitochondria. *The Journal of biological chemistry*, *271*(50), 31763–31766. https://doi.org/10.1074/jbc.271.50.31763

Schernhammer, E. S., & Hankinson, S. E. (2005). Urinary melatonin levels and breast cancer risk. *Journal of the National Cancer Institute*, *97*(14), 1084–1087. https://doi.org/10.1093/jnci/dji190

Schernhammer, E. S., Kroenke, C. H., Laden, F., & Hankinson, S. E. (2006). Night work and risk of breast cancer. *Epidemiology (Cambridge, Mass.)*, *17*(1), 108–111. https://doi.org/10.1097/01.ede.0000190539.03500.c1

Schernhammer, E. S., Laden, F., Speizer, F. E., Willett, W. C., Hunter, D. J., Kawachi, I., & Colditz, G. A. (2001). Rotating night shifts and risk of breast cancer in women participating in the nurses' health study. *Journal of the National Cancer Institute*, *93*(20), 1563–1568. https://doi.org/10.1093/jnci/93.20.1563

Schernhammer, E. S., Rosner, B., Willett, W. C., Laden, F., Colditz, G. A., & Hankinson, S. E. (2004). Epidemiology of urinary melatonin in women and its relation to other hormones and night work. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *13*(6), 936–943.

Schmitz, K. H., Campbell, A. M., Stuiver, M. M., Pinto, B. M., Schwartz, A. L., Morris, G. S., Ligibel, J. A., Cheville, A., Galvão, D. A., Alfano, C. M., Patel, A. V., Hue, T., Gerber, L. H., Sallis, R., Gusani, N. J., Stout, N. L., Chan, L., Flowers, F., Doyle, C., Helmrich, S., … Matthews, C. E. (2019). Exercise is medicine in oncology: Engaging clinicians to help patients move through cancer. *CA: a cancer journal for clinicians, 69*(6), 468–484. https://doi.org/10.3322/caac.21579

Schreiber, G. H., Swaen, G. M., Meijers, J. M., Slangen, J. J., & Sturmans, F. (1993). Cancer mortality and residence near electricity transmission equipment: a retrospective cohort study. *International journal of epidemiology*, *22*(1), 9–15. https://doi.org/10.1093/ije/22.1.9

Seidegård, J., Vorachek, W. R., Pero, R. W., & Pearson, W. R. (1988). Hereditary differences in the expression of the human glutathione transferase active on trans-stilbene oxide are due to a gene deletion. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(19), 7293–7297. https://doi.org/10.1073/pnas.85.19.7293

Seidman, H., Stellman, S. D., & Mushinski, M. H. (1982). A different perspective on breast cancer risk factors: some implications of the nonattributable risk. *CA: a cancer journal for clinicians*, *32*(5), 301–313. https://doi.org/10.3322/canjclin.32.5.301

Seitz, H. K., & Stickel, F. (2007). Molecular mechanisms of alcohol-mediated carcinogenesis. *Nature reviews. Cancer*, *7*(8), 599–612. https://doi.org/10.1038/nrc2191

Shack, L., Jordan, C., Thomson, C. S., Mak, V., Møller, H., & UK Association of Cancer Registries (2008). Variation in incidence of breast, lung and cervical cancer and malignant melanoma of skin by socioeconomic group in England. *BMC cancer*, *8*, 271. https://doi.org/10.1186/1471-2407-8-271

Shangold, M., Freeman, R., Thysen, B., & Gatz, M. (1979). The relationship between long-distance running, plasma progesterone, and luteal phase length. *Fertility and sterility*, *31*(2), 130–133. https://doi.org/10.1016/s0015-0282(16)43811-2

Shekhar, P. V., Werdell, J., & Basrur, V. S. (1997). Environmental estrogen stimulation of growth and estrogen receptor function in preneoplastic and cancerous human breast cell lines. *Journal of the National Cancer Institute*, *89*(23), 1774–1782. https://doi.org/10.1093/jnci/89.23.1774

Shi, L., Remer, T., Buyken, A. E., Hartmann, M. F., Hoffmann, P., & Wudy, S. A. (2010). Prepubertal urinary estrogen excretion and its relationship with pubertal timing. *American journal of physiology. Endocrinology and metabolism*, *299*(6), E990–E997. https://doi.org/10.1152/ajpendo.00374.2010

Shirabe, R., Saito, E., Sawada, N., Ishihara, J., Takachi, R., Abe, S. K., Shimazu, T., Yamaji, T., Goto, A., Iwasaki, M., Inoue, M., Tsugane, S., & JPHC Study Group (2021). Fermented and nonfermented soy foods and the

risk of breast cancer in a Japanese population-based cohort study. *Cancer medicine*, *10*(2), 757–771. https://doi.org/10.1002/cam4.3677

Sigal, A., & Rotter, V. (2000). Oncogenic mutations of the p53 tumor suppressor: the demons of the guardian of the genome. *Cancer research*, *60*(24), 6788–6793.

Simm, P. J., Bajpai, A., Russo, V. C., & Werther, G. A. (2008). Estrogens and growth. *Pediatric endocrinology reviews : PER*, *6*(1), 32–41.

Simpson, P. T., Reis-Filho, J. S., Gale, T., & Lakhani, S. R. (2005). Molecular evolution of breast cancer. *The Journal of pathology*, *205*(2), 248–254. https://doi.org/10.1002/path.1691

Singh, K.K., (1998). Mitochondrial DNA Mutations in Aging, Disease and Cancer. Springer, New York, USA. ISBN:978-3-662-12511-3, 978-3-662-12509-0

Singletary, K. W., Frey, R. S., & Yan, W. (2001). Effect of ethanol on proliferation and estrogen receptor-alpha expression in human breast cancer cells. *Cancer letters*, *165*(2), 131–137. https://doi.org/10.1016/s0304-3835(01)00419-0

Slattery, M. L., Curtin, K., Giuliano, A. R., Sweeney, C., Baumgartner, R., Edwards, S., Wolff, R. K., Baumgartner, K. B., & Byers, T. (2008). Active and passive smoking, IL6, ESR1, and breast cancer risk. *Breast cancer research and treatment*, *109*(1), 101–111. https://doi.org/10.1007/s10549-007-9629-1

Smith-Warner, S. A., Spiegelman, D., Adami, H. O., Beeson, W. L., van den Brandt, P. A., Folsom, A. R., Fraser, G. E., Freudenheim, J. L., Goldbohm, R. A., Graham, S., Kushi, L. H., Miller, A. B., Rohan, T. E., Speizer, F. E., Toniolo, P., Willett, W. C., Wolk, A., Zeleniuch-Jacquotte, A., & Hunter, D. J. (2001). Types of dietary fat and breast cancer: a pooled analysis of cohort studies. *International journal of cancer*, *92*(5), 767–774. https://doi.org/10.1002/1097-0215(20010601)92:5<767::aid-ijc1247>3.0.co;2-0

Smith-Warner, S. A., Spiegelman, D., Yaun, S. S., Adami, H. O., Beeson, W. L., van den Brandt, P. A., Folsom, A. R., Fraser, G. E., Freudenheim, J. L., Goldbohm, R. A., Graham, S., Miller, A. B., Potter, J. D., Rohan, T. E., Speizer, F. E., Toniolo, P., Willett, W. C., Wolk, A., Zeleniuch-Jacquotte, A., & Hunter, D. J. (2001). Intake of fruits and vegetables and risk of breast cancer: a pooled analysis of cohort studies. *JAMA*, *285*(6), 769–776. https://doi.org/10.1001/jama.285.6.769

Smith-Warner, S. A., Spiegelman, D., Yaun, S. S., van den Brandt, P. A., Folsom, A. R., Goldbohm, R. A., Graham, S., Holmberg, L., Howe, G. R., Marshall, J. R., Miller, A. B., Potter, J. D., Speizer, F. E., Willett, W. C., Wolk, A., & Hunter, D. J. (1998). Alcohol and breast cancer in women: a pooled analysis of cohort studies. *JAMA*, *279*(7), 535–540. https://doi.org/10.1001/jama.279.7.535

Society for Investigating the Nature and Cure of Cancer (1806). Edinburgh Medical Surgery Journal, 2:382-9.

Sonnenschein, E., Toniolo, P., Terry, M. B., Bruning, P. F., Kato, I., Koenig, K. L., & Shore, R. E. (1999). Body fat distribution and obesity in pre- and postmenopausal breast cancer. *International journal of epidemiology*, *28*(6), 1026–1031. https://doi.org/10.1093/ije/28.6.1026

Soto, A. M., Sonnenschein, C., Chung, K. L., Fernandez, M. F., Olea, N., & Serrano, F. O. (1995). The E-SCREEN assay as a tool to identify estrogens: an update on estrogenic environmental pollutants. *Environmental health perspectives*, *103 Suppl 7*(Suppl 7), 113–122. https://doi.org/10.1289/ehp.95103s7113

Spink, B. C., Katz, B. H., Hussain, M. M., Pang, S., Connor, S. P., Aldous, K. M., Gierthy, J. F., & Spink, D. C. (2000). SULT1A1 catalyzes 2-methoxyestradiol

sulfonation in MCF-7 breast cancer cells. *Carcinogenesis*, *21*(11), 1947–1957. https://doi.org/10.1093/carcin/21.11.1947

Stallard, S., Litherland, J. C., Cordiner, C. M., Dobson, H. M., George, W. D., Mallon, E. A., & Hole, D. (2000). Effect of hormone replacement therapy on the pathological stage of breast cancer: population based, cross sectional study. *BMJ (Clinical research ed.)*, *320*(7231), 348–349. https://doi.org/10.1136/bmj.320.7231.348

Stevens, R. G., Davis, S., Thomas, D. B., Anderson, L. E., & Wilson, B. W. (1992). Electric power, pineal function, and the risk of breast cancer. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *6*(3), 853–860. https://doi.org/10.1096/fasebj.6.3.1740235

Stevenson, T.H.C. (1915). Annual report of the Registrar-General of births, deaths and marriages in England and Wales. His Majesty's Stationery Office, London, 78:XXXIV.

Strange, R. C., Jones, P. W., & Fryer, A. A. (2000). Glutathione S-transferase: genetics and role in toxicology. *Toxicology letters*, *112-113*, 357–363. https://doi.org/10.1016/s0378-4274(99)00230-1

Sugimura, T. (2000). Nutrition and dietary carcinogens. *Carcinogenesis*, *21*(3), 387–395. https://doi.org/10.1093/carcin/21.3.387

Sun, W. Y., Wu, J. J., Peng, W. T., Sun, J. C., & Wei, W. (2018). The role of G protein-coupled receptor kinases in the pathology of malignant tumors. *Acta pharmacologica Sinica*, *39*(11), 1699–1705. https://doi.org/10.1038/s41401-018-0049-z

Taanman, J. W., (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et biophysica acta*, *1410*(2), 103–123. https://doi.org/10.1016/s0005-2728(98)00161-3

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular biology and evolution, 38(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Tan, D. J., Bai, R. K., & Wong, L. J. (2002). Comprehensive scanning of somatic mitochondrial DNA mutations in breast cancer. *Cancer research*, *62*(4), 972–976.

Tan, D., Goerlitz, D. S., Dumitrescu, R. G., Han, D., Seillier-Moiseiwitsch, F., Spernak, S. M., Orden, R. A., Chen, J., Goldman, R., & Shields, P. G. (2008). Associations between cigarette smoking and mitochondrial DNA abnormalities in buccal cells. *Carcinogenesis*, *29*(6), 1170–1177. https://doi.org/10.1093/carcin/bgn034

Tan, K. L., Webb, G. C., Baker, R. T., & Board, P. G. (1995). Molecular cloning of a cDNA and chromosomal localization of a human theta-class glutathione S-transferase gene (GSTT2) to chromosome 22. *Genomics*, *25*(2), 381–387. https://doi.org/10.1016/0888-7543(95)80037-m

Taningher, M., Malacarne, D., Izzotti, A., Ugolini, D., & Parodi, S. (1999). Drug metabolism polymorphisms as modulators of cancer susceptibility. *Mutation research*, *436*(3), 227–261. https://doi.org/10.1016/s1383-5742(99)00005-8

Tavassoli, F.A., & Devillee, P. (2003). World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of the Breast and Female Organs, IARC Press, Lyon

Tavtigian, S. V., Simard, J., Rommens, J., Couch, F., Shattuck-Eidens, D., Neuhausen, S., Merajver, S., Thorlacius, S., Offit, K., Stoppa-Lyonnet, D., Belanger, C., Bell, R., Berry, S., Bogden, R., Chen, Q., Davis, T., Dumont, M., Frye, C., Hattier, T., Jammulapati, S., Goldgar, D. E. (1996). The complete BRCA2

gene and mutations in chromosome 13q-linked kindreds. *Nature genetics*, *12*(3), 333–337. https://doi.org/10.1038/ng0396-333

Terry, M. B., Zhang, F. F., Kabat, G., Britton, J. A., Teitelbaum, S. L., Neugut, A. I., & Gammon, M. D. (2006). Lifetime alcohol intake and breast cancer risk. *Annals of epidemiology*, *16*(3), 230–240. https://doi.org/10.1016/j.annepidem.2005.06.048

Terry, P. D., & Rohan, T. E. (2002). Cigarette smoking and the risk of breast cancer in women: a review of the literature. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *11*(10 Pt 1), 953–971.

Thangluaia, H. (2012). Mizo kan ni. Mizoram Publication Board, Art & Culture, Government of Mizoram. 81-117.

Tjønneland, A., Christensen, J., Olsen, A., Stripp, C., Thomsen, B. L., Overvad, K., Peeters, P. H., van Gils, C. H., Bueno-de-Mesquita, H. B., Ocké, M. C., Thiebaut, A., Fournier, A., Clavel-Chapelon, F., Berrino, F., Palli, D., Tumino, R., Panico, S., Vineis, P., Agudo, A., Ardanaz, E., … Riboli, E. (2007). Alcohol intake and breast cancer risk: the European Prospective Investigation into Cancer and Nutrition (EPIC). *Cancer causes & control : CCC*, *18*(4), 361–373. https://doi.org/10.1007/s10552-006-0112-9

Tjønneland, A., Thomsen, B. L., Stripp, C., Christensen, J., Overvad, K., Mellemkaer, L., Grønbaek, M., & Olsen, J. H. (2003). Alcohol intake, drinking patterns and risk of postmenopausal breast cancer in Denmark: a prospective cohort study. *Cancer causes & control : CCC*, *14*(3), 277–284. https://doi.org/10.1023/a:1023640720385

Trichopoulos, D., Hsieh, C. C., MacMahon, B., Lin, T. M., Lowe, C. R., Mirra, A. P., Ravnihar, B., Salber, E. J., Valaoras, V. G., & Yuasa, S. (1983). Age at any birth and breast cancer risk. *International journal of cancer*, *31*(6), 701–704. https://doi.org/10.1002/ijc.2910310604

Tseng, L. M., Yin, P. H., Chi, C. W., Hsu, C. Y., Wu, C. W., Lee, L. M., Wei, Y. H., & Lee, H. C. (2006). Mitochondrial DNA mutations and mitochondrial DNA depletion in breast cancer. *Genes, chromosomes & cancer*, *45*(7), 629–638. https://doi.org/10.1002/gcc.20326

Tulinius, H., Day, N. E., Jóhannesson, G., Bjarnason, O., & Gonzales, M. (1978). Reproductive factors and risk for breast cancer in Iceland. *International journal of cancer*, *21*(6), 724–730. https://doi.org/10.1002/ijc.2910210609

Verkasalo, P. K., Lillberg, K., Stevens, R. G., Hublin, C., Partinen, M., Koskenvuo, M., & Kaprio, J. (2005). Sleep duration and breast cancer: a prospective cohort study. *Cancer research*, *65*(20), 9595–9600. https://doi.org/10.1158/0008-5472.CAN-05-2138

Verkasalo, P. K., Thomas, H. V., Appleby, P. N., Davey, G. K., & Key, T. J. (2001). Circulating levels of sex hormones and their relation to risk factors for breast cancer: a cross-sectional study in 1092 pre- and postmenopausal women (United Kingdom). *Cancer causes & control : CCC*, *12*(1), 47–59. https://doi.org/10.1023/a:1008929714862

Verrals, S., (1993). Anatomy and Physiology Applied to Obstetrics. *Embryology, Human*. Churchill Livingstone, Edinburgh, ISBN: 044304211X, 9780443042119. https://www.brainkart.com/media/article/article-The-Breast-Anatomy-Zdh.jpg

Villani, G., Greco, M., Papa, S., & Attardi, G. (1998). Low reserve of cytochrome c oxidase capacity in vivo in the respiratory chain of a variety of human cell

types. *The Journal of biological chemistry*, *273*(48), 31829–31836. https://doi.org/10.1074/jbc.273.48.31829

Voorrips, L. E., Brants, H. A., Kardinaal, A. F., Hiddink, G. J., van den Brandt, P. A., & Goldbohm, R. A. (2002). Intake of conjugated linoleic acid, fat, and other fatty acids in relation to postmenopausal breast cancer: the Netherlands Cohort Study on Diet and Cancer. *The American journal of clinical nutrition*, *76*(4), 873–882. https://doi.org/10.1093/ajcn/76.4.873

Wainwright, J. M. (1931). A Comparison of Conditions Associated with Breast Cancer in Great Britain and America. *The American Journal of Cancer*, 15(4), 2610–2645. https://doi.org/10.1158/ajc.1931.2610

Wang, Y., Liu, V. W., Tsang, P. C., Chiu, P. M., Cheung, A. N., Khoo, U. S., Nagley, P., & Ngan, H. Y. (2006). Microsatellite instability in mitochondrial genome of common female cancers. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society*, *16 Suppl 1*, 259–266. https://doi.org/10.1111/j.1525-1438.2006.00412.x

Warren, M. P. (1980). The effects of exercise on pubertal progression and reproductive function in girls. *The Journal of clinical endocrinology and metabolism*, *51*(5), 1150–1157. https://doi.org/10.1210/jcem-51-5-1150

Wartenberg, D., Calle, E. E., Thun, M. J., Heath, C. W., Jr, Lally, C., & Woodruff, T. (2000). Passive smoking exposure and female breast cancer mortality. *Journal of the National Cancer Institute*, *92*(20), 1666–1673. https://doi.org/10.1093/jnci/92.20.1666

Weiderpass, E., Pukkala, E., Kauppinen, T., Mutanen, P., Paakkulainen, H., Vasama-Neuvonen, K., Boffetta, P., & Partanen, T. (1999). Breast cancer and occupational exposures in women in Finland. *American journal of industrial medicine*, *36*(1), 48–53. https://doi.org/10.1002/(sici)1097-0274(199907)36:1<48::aid-ajim7>3.0.co;2-2

Wellings, S.R. (1980). Development of Human Breast Cancer. *Advances in Cancer Research*, 31, 287-314. https://doi.org/10.1016/S0065-230X(08)60660-0

Welp, E. A., Weiderpass, E., Boffetta, P., Vainio, H., Vasama-Neuvonen, K., Petralia, S., & Partanen, T. J. (1998). Environmental risk factors of breast cancer. *Scandinavian journal of work, environment & health*, *24*(1), 3–7. https://doi.org/10.5271/sjweh.271

Widersten, M., Pearson, W. R., Engström, A., & Mannervik, B. (1991). Heterologous expression of the allelic variant mu-class glutathione transferases mu and psi. *The Biochemical journal*, *276 (Pt 2)*(Pt 2), 519–524. https://doi.org/10.1042/bj2760519

Wiencke, J. K., Kelsey, K. T., Lamela, R. A., & Toscano, W. A., Jr (1990). Human glutathione S-transferase deficiency as a marker of susceptibility to epoxide-induced cytogenetic damage. *Cancer research*, *50*(5), 1585–1590.

Wiencke, J. K., Pemble, S., Ketterer, B., & Kelsey, K. T. (1995). Gene deletion of glutathione S-transferase theta: correlation with induced genetic damage and potential role in endogenous mutagenesis. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *4*(3), 253–259.

Willett, W. C., Browne, M. L., Bain, C., Lipnick, R. J., Stampfer, M. J., Rosner, B., Colditz, G. A., Hennekens, C. H., & Speizer, F. E. (1985). Relative weight and risk of breast cancer among premenopausal women. *American journal of epidemiology*, *122*(5), 731–740. https://doi.org/10.1093/oxfordjournals.aje.a114156

Williams, E. M., Jones, L., Vessey, M. P., & McPherson, K. (1990). Short term increase in risk of breast cancer associated with full term pregnancy. *BMJ*

(Clinical research ed.), *300*(6724), 578–579. https://doi.org/10.1136/bmj.300.6724.578

Winzer, B. M., Whiteman, D. C., Reeves, M. M., & Paratz, J. D. (2011). Physical activity and cancer prevention: a systematic review of clinical trials. *Cancer causes & control: CCC*, *22*(6), 811–826. https://doi.org/10.1007/s10552-011-9761-4

Wiseman, M. (2008). The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *The Proceedings of the Nutrition Society*, *67*(3), 253–256. https://doi.org/10.1017/S002966510800712X

Wolf, I., Sadetzki, S., Catane, R., Karasik, A., & Kaufman, B. (2005). Diabetes mellitus and breast cancer. *The Lancet. Oncology*, *6*(2), 103–111. https://doi.org/10.1016/S1470-2045(05)01736-5

Wolk, A., Bergström, R., Hunter, D., Willett, W., Ljung, H., Holmberg, L., Bergkvist, L., Bruce, A., & Adami, H. O. (1998). A prospective study of association of monounsaturated fat and other types of fat with risk of breast cancer. *Archives of internal medicine*, *158*(1), 41–45. https://doi.org/10.1001/archinte.158.1.41

Woods, K. L., Smith, S. R., & Morrison, J. M. (1980). Parity and breast cancer: evidence of a dual effect. *British medical journal*, *281*(6237), 419–421. https://doi.org/10.1136/bmj.281.6237.419

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., & Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature*, *378*(6559), 789–792. https://doi.org/10.1038/378789a0

Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., & Averill, D. (1994). Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science (New York, N.Y.)*, *265*(5181), 2088–2090. https://doi.org/10.1126/science.8091231

World Cancer Research Fund, & American Institute for Cancer Research. (2007). Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *Nutrition (Burbank, Los Angeles), Vol. 1.*

World Cancer Research Fund, & American Institute for Cancer Research. (2010). Continuous update project report. Food, nutrition, physical activity, and the prevention of breast cancer. American Institute for Cancer Research.

Wrede, J. E., Mengel-From, J., Buchwald, D., Vitiello, M. V., Bamshad, M., Noonan, C., Christiansen, L., Christensen, K., & Watson, N. F. (2015). Mitochondrial DNA Copy Number in Sleep Duration Discordant Monozygotic Twins. *Sleep*, *38*(10), 1655–1658. https://doi.org/10.5665/sleep.5068

Wright, R. M., McManaman, J. L., & Repine, J. E. (1999). Alcohol-induced breast cancer: a proposed mechanism. *Free radical biology & medicine*, *26*(3-4), 348–354. https://doi.org/10.1016/s0891-5849(98)00204-4

Wu, A. H., Pike, M. C., & Stram, D. O. (1999). Meta-analysis: dietary fat intake, serum estrogen levels, and the risk of breast cancer. *Journal of the National Cancer Institute*, *91*(6), 529–534. https://doi.org/10.1093/jnci/91.6.529

Wu, A. H., Yu, M. C., Tseng, C. C., Stanczyk, F. Z., & Pike, M. C. (2007). Diabetes and risk of breast cancer in Asian-American women. *Carcinogenesis*, *28*(7), 1561–1566. https://doi.org/10.1093/carcin/bgm081

Wu, B., & Dong, D. (2012). Human cytosolic glutathione transferases: structure, function, and drug discovery. *Trends in pharmacological sciences*, *33*(12), 656–668. https://doi.org/10.1016/j.tips.2012.09.007

Xu, C., Tran-Thanh, D., Ma, C., May, K., Jung, J., Vecchiarelli, J., & Done, S. J. (2012). Mitochondrial D310 mutations in the early development of breast cancer. *British journal of cancer*, *106*(9), 1506–1511. https://doi.org/10.1038/bjc.2012.74

Xue, F., & Michels, K. B. (2007). Diabetes, metabolic syndrome, and breast cancer: a review of the current evidence. *The American journal of clinical nutrition*, *86*(3), s823–s835. https://doi.org/10.1093/ajcn/86.3.823S

Yager, J. D., & Davidson, N. E. (2006). Estrogen carcinogenesis in breast cancer. *The New England journal of medicine*, *354*(3), 270–282. https://doi.org/10.1056/NEJMra050776

Yang, L., & Jacobsen, K. H. (2008). A systematic review of the association between breastfeeding and breast cancer. *Journal of women's health (2002)*, *17*(10), 1635–1645. https://doi.org/10.1089/jwh.2008.0917

Ye, C., Gao, Y. T., Wen, W., Breyer, J. P., Shu, X. O., Smith, J. R., Zheng, W., & Cai, Q. (2008). Association of mitochondrial DNA displacement loop (CA)n dinucleotide repeat polymorphism with breast cancer risk and survival among Chinese women. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, *17*(8), 2117–2122. https://doi.org/10.1158/1055-9965.EPI-07-2798

Ye, C., Shu, X. O., Pierce, L., Wen, W., Courtney, R., Gao, Y. T., Zheng, W., & Cai, Q. (2010). Mutations in the mitochondrial DNA D-loop region and breast cancer risk. *Breast cancer research and treatment*, *119*(2), 431–436. https://doi.org/10.1007/s10549-009-0397-y

Yoshimura, A., Imoto, I., & Iwata, H. (2022). Functions of Breast Cancer Predisposition Genes: Implications for Clinical Management. *International journal of molecular sciences*, *23*(13), 7481. https://doi.org/10.3390/ijms23137481

Yu, M., Shi, Y., Zhang, F., Zhou, Y., Yang, Y., Wei, X., Zhang, L., & Niu, R. (2008). Sequence variations of mitochondrial DNA D-loop region are highly frequent events in familial breast cancer. *Journal of biomedical science*, *15*(4), 535–543. https://doi.org/10.1007/s11373-007-9229-4

Yu, Z. G., Jia, C. X., Geng, C. Z., Tang, J. H., Zhang, J., & Liu, L. Y. (2012). Risk factors related to female breast cancer in regions of Northeast China: a 1:3 matched case-control population-based study. *Chinese medical journal*, *125*(5), 733–740.

Yue, W., Wang, J., Schiermeyer, K., Xu, P., Gil, J., & Felder, R. (2019). Role of GRK4, a Risk Factor for Hypertension in Breast Cancer. *Hypertension*, 74:AP2043. https://doi.org/10.1161/hyp.74.suppl_1.P2043

Zheng, W., Gustafson, D. R., Sinha, R., Cerhan, J. R., Moore, D., Hong, C. P., Anderson, K. E., Kushi, L. H., Sellers, T. A., & Folsom, A. R. (1998). Well-done meat intake and the risk of breast cancer. *Journal of the National Cancer Institute*, *90*(22), 1724–1729. https://doi.org/10.1093/jnci/90.22.1724

Zhong, S., Wyllie, A. H., Barnes, D., Wolf, C. R., & Spurr, N. K. (1993). Relationship between the GSTM1 genetic polymorphism and susceptibility to bladder, breast and colon cancer. *Carcinogenesis*, *14*(9), 1821–1824. https://doi.org/10.1093/carcin/14.9.1821

Ziegler, R. G., Hoover, R. N., Pike, M. C., Hildesheim, A., Nomura, A. M., West, D. W., Wu-Williams, A. H., Kolonel, L. N., Horn-Ross, P. L., Rosenthal, J. F., & Hyer, M. B. (1993). Migration patterns and breast cancer risk in Asian-American women. *Journal of the National Cancer Institute*, *85*(22), 1819–1827. https://doi.org/10.1093/jnci/85.22.1819

BRIEF BIO-DATA OF THE CANDIDATE

Name : **Lalhmachhuani**

Designation : **Research Scholar**

Postal Address : B-58, Zorun

Mac Donald Hill, Zarkawt

Aizawl -796007, Mizoram

Mobile No. : 9436141876

E-mail : chhuani17@gmail.com

Date of Birth : January 20, 1975

Educational Qualification :

Degree	Institution	Marks (%)	Year
B. Sc (Zoology)	North Eastern Hill University	52.8	1996
M.Sc. (Zoology)	Madras University	62.5	1998

Research Experience :

Conferences/Presentation :

Other Information, if any : NIL

PARTICULARS OF THE CANDIDATE

NAME OF THE CANDIDATE	:	**LALHMACHHUANI**
DEGREE	:	**DOCTOR OF PHILOSOPHY**
DEPARTMENT	:	**ZOOLOGY**
TITLE OF THE THESIS	:	**Mutational analysis of mitochondrial and glutathione s-transferase genes associated with breast cancer among Mizo population**
DATE OF ADMISSION	:	August 8, 2013
APPROVAL OF RESEARCH PROPOSAL		
1. BOS	:	May 16, 2014
2. SCHOOL BOARD	:	May 23, 2014
MZU REGISTRATION NO.	:	**235 of 2014**
PH.D. REGISTRATION NO. & DATE	:	**MZU/Ph.D./635 of 23.05.2014**
EXTENSION (IF ANY)	:	**No.16-2/MZU(Acad)/21/313-318 of 12.7.23**

HEAD

Department of Zoology

ABSTRACT

MUTATIONAL ANALYSIS OF MITOCHONDRIAL AND GLUTATHIONE S-TRANSFERASE GENES ASSOCIATED WITH BREAST CANCER RISK AMONG MIZO POPULATION

AN ABSTRACT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

LALHMACHHUANI

MZU REGISTRATION: 235 of 2014

Ph.D. REGISTRATION: MZU/Ph.D./635 of 23.05.2014

DEPARTMENT OF ZOOLOGY

SCHOOL OF LIFE SCIENCES

JUNE, 2023

MUTATIONAL ANALYSIS OF MITOCHONDRIAL AND GLUTATHIONE S-TRANSFERASE GENES ASSOCIATED WITH BREAST CANCER RISK AMONG MIZO POPULATION

BY

Lalhmachhuani

Department of Zoology

Prof. G. Gurusubramanian

Submitted

In partial fulfillment of the requirement of the Degree of

Doctor of Philosophy in Zoology of Mizoram University, Aizawl

1. Introduction

The Mizos are a group of tribes inhabiting a landlocked area located on the north eastern part of India called Mizoram. It has an area of 21,081 square kilometers located at 21°56'N to 24°31'N latitude and 92°16'E to 93°26'E longitude. Mizoram shares its boundary on the northern side with Tripura, Assam and Manipur; and an international boundary with Myanmar in the east and south, and Bangladesh in the west. The Mizos are different from mainland India in their race and ethnicity and have a unique lifestyle, customs and dietary habits.

There is a saying that many women who develop breast cancer have no known risk factors other than simply being women. But this alone does not explain the 'why', the randomness and the occasional incidence in men. We have come a long way from the time when breast cancer was first described in the Edwin Smith Papyrus, an ancient Egyptian medical text as an 'ulcer' and untreatable (Diamandopoulos, 1996). We know now that there is a hereditary component to breast cancer and that it tends to cluster in certain families. Studies of migrants from low-risk Asian population migrating to high-risk USA provided the first solid evidence that environmental determinants were responsible for most of the observed international and inter-ethnic differences in breast cancer incidence (Ziegler *et al.*, 1993). The relative contributions of pure genetic effects and of lifestyle remain unclear, but it is becoming more and more evident that genotypic inheritance and lifestyle are probably inseparably intertwined. The reason could be because the combination of genetic factors and lifestyle makes us who we are, and determining our individual susceptibility to that disease (Key *et al.*, 2001).

Breast cancer incidence rates are high in more developed countries, whereas rates in less developed countries are low but increasing. The volume on breast cancer research is huge, but there is no proper consensus on their findings, which could be due to differences in sample sizes, geography, race and ethnicity of study populations, local customs, lifestyle, culture, and health care conditions (Parkin *et al.*, 2005). Due to

variation in the geographical, racial and ethnic distribution of the disease, it is of utmost importance to identify risk factors that is specific for that population and whether the established factors in other parts of the world applicable or relevant for that particular region. The purpose of the investigations in this thesis is to cast some light on which recognized environmental and lifestyle factors could be responsible for breast cancer incidence among the Mizo population. The other objective is to determine whether mutations in the mitochondrial and glutathione s-transferases genes might interact with these environmental and lifestyle factors in increasing breast cancer risk.

1.1. Biology of the Breast

The basic component of a breast is alveoli made up of epithelial cuboidal cells embedded in myoepithelial cells. Each alveolus is lined by milk-secreting cells, the acini, which extract from the mammary blood supply the factors essential for milk formation. Each breast has 15 to 20 sections called lobes and within each lobe are many smaller lobules. Lobules are arranged in clusters like a bunch of grapes. Lobules unite together through a mesh work of ducts called lactiferous ducts. Lobules end in dozens of tiny bulbs that can produce milk. Thin tubes, called ducts, link all the lobes, lobules and bulbs. These ducts lead to the nipple in the center of a dark area of skin called the areola. Fat fills the spaces between lobules and ducts. There are no muscles in the breast, but muscles lie under each breast and cover the ribs. Branching of the ducts and alveolar growth occurs before puberty due to ovarian estrogen stimulation. Lobule formation in the female breast occurs within 1-2 years after the onset of the first menstrual period. Full differentiation of the mammary gland is a gradual process, which takes years to complete. It has also been observed that full maturation of secretory alveoli occurs at pregnancy, in which increased estrogen and progesterone levels cause further differentiation of ductal cells (Imagawa *et al*., 1990).

The breast of nulliparous women contains more undifferentiated structures while in premenopausal parous women the predominant structure is the most

differentiated form. Full lobular differentiation only occurs in parous women, especially in those experiencing full term pregnancy at a young age. After menopause, the breast undergoes regression in both nulliparous and parous women manifested as an increase in the undifferentiated structures. Estrogens are a group of hormones that play a major role in promoting the proliferation of both normal and neoplastic breast epithelium. The amounts and types of estrogen present vary throughout life. The two types of estrogen that have been most closely studied in relation to breast cancer risk are estradiol and estrone sulfate. Estrone is the primary form after menopause which is produced mainly by fat cells, estradiol is the primary form during reproductive years made in ovaries, and estriol is the primary form during pregnancy. Estradiol acts locally on the mammary gland, stimulating DNA synthesis and promoting bud formation (Russo et al., 2000).

1.2. Types of Breast Cancer

All tumors arise from normal tissue and breast cancer is cancer that originates from breast tissue but the progression from normal breast tissue to invasive cancer is poorly understood. The most common type of breast cancer begins in the lining of the ducts called ductal carcinoma. The second most common is lobular carcinoma, which occurs in the lobes. Non-invasive breast cancer is called carcinoma in situ (CIS) and can arise from either ductal or lobular hyperplasia of epithelial cells (Wellings, 1980). Cancer that has progressed into surrounding tissue is called invasive breast cancer and usually has the ability to metastasize. Tumors are categorized according to type and size, histopathology, invasiveness, tumor stage and receptor expression. With improvement in molecular techniques, we now have a deeper understanding of diverse breast cancer types and how they differ (Simpson *et al*., 2005).

According to WHO, tumors are classified into six main types - ductal, lobular, mucinous, medullary, papillary and tubular carcinoma. Ductal and lobular tumors represent around 90-95% of all cases (Tavassoli and Devilee, 2003). Histological

grade is often classified according to the Nottingham Grade classification which was introduced in the 1990s and includes three different parameters - tubule formation, nuclear pleomorphism and mitotic counts (Elston and Ellis, 1991). Tumor stage classification incorporates Tumor size (T), lymph Node status (N) and Metastasis (M) (usually shortened to TNM). The TNM system has been somewhat controversial but remains well used by clinicians (Benson *et al*., 2003; Cserni *et al*., 2018). Expression of different receptors, known to affect the prognostic and predictive values of therapy, is also used to characterize the tumors. They are classified according to expression of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2). The lack of any ER or PR on the tumor cells makes the prognosis worse whereas the lack of HER2 expression does not. If the tumor lacks all three receptors it is called triple negative, this feature often indicates a poor prognosis (Bauer *et al*., 2007; Parise *et al*., 2009).

1.3. Epidemiology

Risk factor is anything that affects the chance of getting breast cancer. A statement by Jose Russo sums up the most appropriate definition of one's risk saying "Breast cancer is a complex disease caused by multiple environmental and lifestyle factors interacting with genetic susceptibility across the life span." There are important findings that establish certain agents as risk factors. These studies have shown that there are modifiable and non-modifiable factors that influence breast cancer risk. There are several controversial risk factors which are uncertain but suspected to cause breast cancer as well. Multiple occurrences of these risk factors in a person increase the likelihood of developing breast cancer.

According to WHO, breast cancer is the most frequent cancer among women, impacting 2.1 million women each year, and among women the cause for the greatest number of cancer-related deaths. According to GLOBOCAN 2012, among women, 25.2 percent of incident sites of cancer was the breast and had a substantially higher

incidence (43.3 per 100 000) than any other cancer; the next highest incidence was of colorectal cancer (14.3 per 100 000). In 2018, it was estimated that 15 percent of all cancer deaths among women was due to breast cancer. Breast cancer rates used to be higher among women in more developed regions, but according to recent statistics, it is increasing globally in nearly every region. In India, roughly, one in four newly detected cancers in women is breast cancer. In 2018, 1,62,468 new cases and 87,090 deaths due to breast cancer were reported. The global distribution of cancer indicates marked, and sometimes extreme differences with respect to particular tumor type, which could be the key to understanding causation, and hence the development of preventive measures. In Mizoram, according to PBCR 2012-2014 report, breast cancer is the third most common cancer among females (13.5%), but we see rise in incidence and could lead to the most common cancer among females.

1.4. Mitochondrial DNA

Mitochondria play a central role in oxidative metabolism and each cell contains thousands of copies of mitochondrial DNA (mtDNA). The human mtDNA genome is a double stranded 16.6 kb circular DNA and is formed by a light strand and a heavy strand. Protein coding segments on mtDNA do not have introns and are transcribed by a single polycistronic mRNA from each strand. Mutations in the mtDNA have been found in connection to various types of human cancer. Since the mtDNA encode several polypeptides of the respiratory-chain enzymes, mtDNA mutations often affect the function of oxidative phosphorylation. Some of the identified mutations cause amino-acid substitutions in the enzyme cytochrome c oxidase. Mitochondrial DNA is particularly susceptible to damage by environmental carcinogens because it contains no introns and lack histones, and is in close proximity to reactive oxygen species (ROS) produced through oxidative phosphorylation. Due to this, the mutation frequency in mtDNA is approximately tenfold greater than that in nuclear DNA (Johns, 1995; Grossman and Shoubridge, 1996).

1.4.1. Displacement Loop

The d-loop is a triple-stranded non-coding region of mtDNA, 1124 bp in size (np 16024-516) that contains cis-regulatory elements required for the replication and transcription of the mtDNA. All other mitochondrial proteins, including those involved in the replication, transcription and translation of mtDNA are nuclear encoded (Schatz, 1996). The d-loop region contains the leading strand for the origin of replication and a number of major promoters for transcription of the mitochondrial genome. Hence, it is possible that genetic variability in the d-loop region may affect the function of the respiration chain that is responsible for high reactive oxygen species levels and could contribute to nuclear genome damage and cancer initiation and progression. Moreover, respiratory chain alteration may cause a dysfunction in mitochondrion induced apoptosis (Ye *et al*., 2010).

1.4.2. Cytochrome C Oxidase

Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. It is the rate-limiting step of the mitochondrial electron transport chain (Villani *et al*., 1998) and represents a molecular switch that induces apoptosis under energy stress conditions (Kadenbach *et al*., 2004). Defects in the mitochondrial genome and function are suspected to contribute to the development and progression of cancer (Ye *et al*., 2008). Several of the cancer associated mutations found in mtDNA result in structural modifications of cytochrome c oxidase (Namslauer and Brzezinksi, 2009). Of the 13 subunits that comprise cytochrome c oxidase, subunit I (np 5904-7445) is responsible for the control of apoptosis through phosphorylation or dephosphorylation events (Lee *et al*., 2001).

1.5. Glutathione S-transferase

The glutathione s-transferases (GSTs) are complex superfamily of dimeric phase II metabolic enzymes involved in detoxification of a wide range of harmful

chemicals, including environmental pollutants, carcinogens, mutagens and toxic products such as lipid hydrogen peroxides generated during oxidative stress. A single GST unit consists of two subunits an N-terminal α/β-domain or G domain for binding glutathione (GSH) and an all-α-helical domain or H domain for binding hydrophobic substrates. The residues forming the glutathione binding site are conserved in the different classes, while those forming the substrate binding site vary considerably, leading to a wide substrate specificity (Johansson and Mannervik, 2001). Human GSTs consist of three families - cytosolic GSTs, mitochondrial GSTs and microsomal GSTs (Hayes *et al*., 2005). Cytosolic GSTs are further categorized into seven major classes - alpha (five members), mu (five members), pi (one member), theta (two members), zeta (one member), omega (two members), and sigma (one member). Members of the same class possess greater than 40 percent amino acid sequence identity and between classes, proteins have less than 25 percent sequence identity (Mannervik *et al*., 2005).

GSTs play an important role in cellular defense as they catalyze the conjugation of reduced glutathione with various electrophilic compounds such as the one present in tobacco smoke (Ishii *et al*., 1999). Associations between GST genotypes and disease phenotype may reflect a link between alleles and cytogenetic damage and specific mutations in target genes. GSTM1 and GSTT1 null cells are more susceptible to sister chromatid exchange following exposure to various electrophiles. Presumably genotypes, alone or in combination, should identify subjects who are detoxification deficient and consequently more likely to suffer formation of carcinogen DNA adducts and/or mutations (Ryberg *et al*., 1997). Thus, these two loci, GSTM1 mapped on chromosome 1p13.3 and GSTT1 on chromosome 22q11.23 are studied in particular because of their relevance in indicating susceptibility to cancer. Studies have shown that individuals who inherit the GSTM1 null genotype are not capable of conjugating and detoxifying specific substrate epoxide intermediates (Wiencke *et al*., 1990). Thus, the absence of the GSTM1 gene should increase cancer risk from environmental

exposure while the presence of the intact GSTM1 gene would be protective for cytogenetic damage and carcinogen derived DNA adduct formation. GSTT1 has also been involved in the glutathione-dependent detoxification. Similar to GSTM1, GSTT1 has significant activity towards epoxides, suggesting that individuals without both GSTM1 and GSTT1 may be at a particularly high risk of cancer (Wiencke *et al*., 1995).

In conclusion, in spite of numerous published data on breast cancer etiology, there is no proper consensus on these findings. This could be due to variation in the geographical, racial and ethnic distribution of the disease. There are also very few studies to confirm that the risk factors established in other parts of the world be the reason for incidence among the Mizo population. With rise in breast cancer incidence in this region, it is of great importance to identify the risk factors specific for this region.

2. Review of Literature

2.1. Brief History of Breast Cancer

The first authentic accounts of breast cancer can be traced back to 3,000-2,500 B.C., in ancient Egypt (Breasted, 1930). In 400 B.C., Hippocrates described the progressive stages of breast cancer and associated the origin of breast cancer with cessation of menstruation (Ariel, 1987). In 200 A.D., Galen attributed the accumulation of black bile in the blood to cause breast cancer (De Moulin, 1983). In 1713, Bernardino Ramazzini in Italy, noted a higher frequency of breast cancer in nuns than in married women (Pope, 2004). In 1806, the Society for Investigating the Nature and Cure of Cancer published the findings of a questionnaire about the disease commenting, 'with regard to cancer, it is not only necessary to observe the effects of climate and local situation but to extend our views to different employments, as those in various metals and manufactures; in mines and collieries; in the army and navy; in those who lead sedentary or active lives; in the married or single; in the different sexes, and many other circumstances. Should it be proved that women are more subject to cancer than men, then we may enquire whether married women are more liable to have the uterus or breast affected; those who have suckled or those who did not; and the same observations may be made of the single' (Society for Investigating the Nature and Cure of Cancer, 1806). In 1842, Rigoni Stern in Italy, compared the incidence of cancer of the breast and uterus among married and unmarried females and showed the relationship of marital status to these cancers (Rigoni-Stern, 1842).

In 1915, a major contribution to breast cancer epidemiology came from an annual report of the Registrar-General of births, deaths and marriages in England and Wales. This report indicated that the mortality rate for breast cancer was markedly higher in single than in married women after the age of 45 (Stevenson, 1915). In 1926, Janet Elizabeth Lane-Claypon carried out the first modern case-control study in United Kingdom comparing 508 breast cancer patients with 509 healthy women. This

study concluded that low fertility increases breast cancer risk (Lane-Claypon, 1926). In 1931, Lane-Claypon's study was replicated by JM Wainwright using a United States sample of 679 breast cancer cases and 567 unmatched controls. The 1926 United Kingdom study and 1931 United States study marked the beginning of a new era of etiologic research as it provided the first evidence from observational studies that parity, age at marriage, and artificial menopause were associated with breast cancer risk (Wainwright, 1931).

In 1866, Paul Broca wrote about the high prevalence of breast carcinoma in his wife's family and identified cases up to four generations from her family (Broca, 1866). This is the first of many reports that pointed out heritability of breast cancer and increased susceptibility for persons having positive family history of breast cancer (Claus *et al*., 1998). Key evidence that a single genetic mutation could cause heritable breast cancer risk came with the identification of a locus on chromosome 17q that was linked to disease susceptibility in specific families (Hall *et al*., 1990). In 1994, the BRCA1 gene was subsequently identified through positional cloning (Miki *et al*., 1994). During the same year, the second breast cancer susceptibility locus was localized to chromosome 13q12-13 by linkage studies of families with multiple cases of early-onset breast cancer that were not linked to BRCA1 (Wooster *et al*., 1994). The BRCA2 gene was cloned in 1995 (Wooster *et al*., 1995) and its complete coding sequence and exonic structure were described in 1996 (Tavtigian *et al*., 1996). ATM, BARD1, CHEK2, PALB2, TP53, PTEN, CDH1, STK11, RAD51C and RAD51D have also been recognized as breast cancer predisposition genes with a high to moderate risk. Most breast cancer predisposition genes participate in DNA damage repair pathways and cell cycle checkpoint mechanisms (Yoshimura *et al*., 2022).

The molecular mechanisms underlying the development of breast cancer are not completely understood. These tumours are likely to be caused by the interaction between many genetic and environmental factors. It was found that chemicals,

radiation and viruses inflict potential harm at certain regions on the hereditary material and their presence leads to impairment in the functionality of several genes (Cook *et al*., 1933; Riou *et al*., 1990). It was also observed that certain changes termed as germline mutations on this macromolecule were heritable. This further elicits the mechanistic approach regarding a better understanding of disease progression (Malik *et al*., 2009). With the advent of modern technologies, researchers and oncologists had a better understanding of cancer not only at the cellular but also at the macro and micro molecular levels. It is becoming more evident that in most cases, genotypic inheritance and lifestyle are probably inseparably intertwined. It is generally believed that the initiation of breast cancer is a consequence of cumulative genetic damages leading to genetic alterations resulting in activation of proto-oncogenes and inactivation of tumor suppressor genes, followed by uncontrolled cellular proliferation and/or aberrant programmed cell death or apoptosis. The role of reactive oxygen species (ROS) has been related to the etiology of cancer (Emerit, 1994).

2.2. Epidemiology

The statement that one out of every eight women will develop breast cancer in their lifetime indicates the gravity of this disease (Ferrini *et al*., 2015). Even though breast cancer account for the most number of all new cases of cancer in females, it ranks as the fifth most common cause of death, because of the relatively more favourable prognosis (mortality to incidence ratio, 0.35) making it the most prevalent cancer in the world today (Parkin, 2004). The global burden of breast cancer continues to rise with over one million new cases diagnosed and 400,000 deaths occurring each year in women. According to GLOBOCAN an online database that maintains cancer statistics for 185 countries, in 2020 there were 78 lakhs women alive who were diagnosed with breast cancer in the past 5 years; and 26 percent of all newly diagnosed cancer is cancer of the breast. In India, breast cancer accounted for 13.5 percent of all cancer cases and 10.6 percent of all deaths.

According to PBCR 2012-2014 report, the highest age adjusted rate (AAR) of breast cancer were in Delhi (41.0), Chennai (37.9), Bangalore (34.4) and Thiruvananthapuram District (33.7). In Mizoram, breast cancer is at a crude rate (CR) of 15.8 and AAR of 19.9 per 100,000 population, with breast cancer being the third (13.5%) most common cancer among females after cervix uteri (15.4%) and lung (14.1%). In just Aizawl district, breast cancer is third (14.5%) most common after lung (17.8%) and cervix uteri (15.6%). In the other seven districts excluding Aizawl, breast cancer is the fourth (11.3%) most common after cervix uteri (16.3%), lung (13.3%) and stomach (12.1%). According to National Cancer Registry Programme (NCRP) 2020 report, the highest breast cancer incidence among all the North-Eastern PBCRs with AAR 30.7 is Aizawl district (Mathur *et al*., 2020).

Breast cancer incidence rate varies widely worldwide due to a range of socio-economic, reproductive, hormonal, nutritional and genetic factors which can be broadly classified into two factors, modifiable and non-modifiable (McPherson *et al*., 2000). Incidence is more closely associated to age than to any other risk factor, it increases rapidly during the reproductive years and then more slowly after about 50 years of age (Key *et al*., 2001). About 55 percent of the global burden is among developed countries, but incidence rates are rapidly rising in developing countries. While incidence rate is less than 40 per 100,000 women in most of the less developed countries, breast cancer is still the most common cancer among women in these countries (Parkin *et al*., 2005). A history of benign breast disease is also a well-established risk factor for breast cancer. Women with severe atypical epithelial hyperplasia have been found to have significantly higher risk compared to women who do not have any proliferative changes in their breast (McPherson *et al*., 2000). Numerous studies have been conducted to find the risk factors for breast cancer, some of the major recognized and suspected factors are listed as follows -

2.2.1. Reproductive Factors

The role of reproductive factors must be one of the most studied factors in breast cancer risk. Studies have revealed that a woman's breast undergoes many changes especially from puberty till menopause exposing it to high endogenous hormone levels (Key, 1999; Kabuto *et al*., 2000). Steroid hormones stimulate cellular replication and mitotic activity in breast epithelium which are believed to be crucial in the pathogenesis of mammary cancer. This high rate of cell division increases the frequency and likelihood of propagation of copying errors and DNA changes (Pike *et al*., 1983). Results from animal studies indicate that estrogen metabolites have genotoxic properties (Yager and Davidson, 2006). Lifetime exposure to endogenous sex hormones is determined by several variables including timing of menarche, age at first full term pregnancy, number of pregnancies, and age at menopause (Feigelson and Henderson, 1996). Cumulative lifetime exposure to estrogen is thought to be a key factor in determining breast cancer risk in women (Henderson *et al*., 1985).

2.2.1.1. Age at Menarche and Menopause

The milestone events that determine the period over which women are exposed to endogenous ovarian hormones have repeatedly been reported to influence breast cancer risk (Tulinius *et al*., 1978; Kvåle and Heuch, 1988; Hsieh *et al*., 1990). Late onsets of menarche as well as early menopause are associated with significant decrease in risk of 5 percent per year and 3 percent per year respectively (Collaborative Group on Hormonal Factors in Breast Cancer, 1997). The increased breast cancer risk associated with early age at menarche (< 12 years) (Kelsey *et al*., 1993; Bernstein *et al*., 1994; Berkey *et al*., 1999; Maurya and Brahmachari, 2022) is probably due to prolonged exposure of breast epithelium to estrogen with earlier onset of regular menstrual cycles and higher estrogen levels for longer years (Henderson *et al*., 1985; Apter *et al*., 1989). Urinary estrogens are significantly higher in girls who have early menarche than in those with normal menarche (Shi *et al*., 2010). Similarly, later age at menopause

maximises the number of ovulatory cycles and may therefore lead to increased risk. For every one year increase in the age at menopause, the risk of breast cancer increases by approximately 3 percent (Collaborative Group on Hormonal Factors in Breast Cancer, 1996). One of the most compelling pieces of evidence regarding the influence of endogenous hormones on breast cancer risk is found in the levelling off in the age-specific incidence curve of breast cancer after menopause when ovarian production of steroid hormones ceases. Breast cancer risk in women with natural menopause before the age of 45 years is half compared with women who stop menstruating after the age of 55 years (Pike *et al*., 2004). There is also convincing evidence of an age-dependent protective effect of early menopause surgically induced by bilateral oophorectomy (Brinton *et al*., 1988; Li *et al*., 2016).

2.2.1.2. Childbearing

The exact mechanism by which an early first birth protects against breast cancer remains incompletely understood, but has primarily been attributed to shortening of the time window of high susceptibility to carcinogenic transformations (Russo *et al*., 2000). A woman's risk of breast cancer appears to be related to timing of first birth and age at subsequent births. Higher parity and early age at first birth have both been associated with decreased lifetime incidence of breast cancer (Kelsey *et al*., 1993; Chie *et al*., 2000; Pathak *et al*., 2000). Studies in India among Parsi women, who on average are wealthy, marry late, and have few children, have an age-adjusted incidence rate of breast cancer that is more than twice that of Hindu women living in the same geographical area, who as a group are poorer, marries earlier, and have more children (Jussawalla *et al*., 1981).

In most Western countries, there is a social gradient in breast cancer risk with markedly higher incidence in women with high education compared to women with low education (Faggiano *et al*., 1997; Shack *et al*., 2008). For a woman to have higher education, childbearing is usually delayed in most cases. Breast cancer risk is half in

women having their first child before the age of 20 compared to those having their first child after the age of 30 (Russo *et al*., 2000). There is also some evidence that the interval between age at menarche and age at first birth may be relevant to breast cancer risk (Andrieu *et al*., 2000; Clavel-Chapelon and E3N Group, 2002; Li *et al*., 2008).

A study in 1970 observed protective effect of parity attributing to an earlier age at first childbirth in women with many children (MacMahon *et al*., 1970). It is now estimated that for each additional year of age at first birth, the risk of premenopausal breast cancer increases by 5 percent, and increases by 3 percent for breast cancers diagnosed after menopause (Clavel-Chapelon and Gerber, 2002). Compared to nulliparous women, women with a first full-term pregnancy before age 20 years have about half the risk of breast cancer (Kelsey *et al*., 1993). Women with an older age at first birth (≥35 years) have the same risk of breast cancer as nulliparous women. In a reanalysis of MacMahon's data, older age at any birth was found to be an independent risk indicator (Trichopoulos *et al*., 1983).

2.2.1.3. Parity

It has long been recognized that parity reduces the risk of breast cancer (Kelsey *et al*., 1993). In 1713, Ramazzini of Padua observed what appeared to be an epidemic of breast cancer among nuns (Pope, 2004). One hundred years later, it was noted that breast cancer was at least three times as frequent in nuns as in other women (Rigoni-Stern, 1842). The risk reduction appears to be greatest among women with high parity, where the risk reduction due to breastfeeding may be as great as 50 percent (Romieu *et al*., 1996) and among premenopausal women with lactation durations ≥2 years, where the breast cancer risk reduction may be 30 percent (Newcomb *et al*., 1994). There is also evidence that the timing of pregnancy is relevant to breast cancer risk. Compared to single women, the risk of breast cancer is lower in older married women, but not in younger married women, with an approximate crossover of the effect around age 40 (Janerich and Hoff, 1982; Pathak *et al*., 1986). There are other reports of a higher

breast cancer risk among young parous compared to young nulliparous women (Woods *et al*., 1980; Layde *et al*., 1989) and an increased risk of breast cancer in the years following childbirth (Bruzzi *et al*., 1988; Layde *et al*., 1989; Williams *et al*., 1990). Pregnancy induces both transient and permanent structural changes in the breast tissue of laboratory animals (Russo *et al*., 1982, 1990). It appears that the effect of parity is determined by the age of occurrence of component pregnancies and that the closer the births are together, the lower the risk. A likely explanation is that pregnancies occurring close together in time provide less time for breast cells to accumulate DNA damage and that every new pregnancy affords additional protection by recruiting more of the remaining undifferentiated cells (Russo and Russo, 1993).

There is much debate whether an incomplete pregnancy affects future breast cancer risk. Based on findings from animal studies, it has been hypothesized that an increase in breast cancer risk may follow if the hormonal surge occurring during the first trimester is not followed by the protective components of breast tissue maturation and terminal differentiation of lobular structures during the second and third trimester (Russo *et al*., 1982). Findings from case control studies indicated that induced abortions were associated with an increased risk of breast cancer (Michels and Willett, 1996). However, collective evidence to date points to no association between pregnancy interruption and subsequent breast cancer risk. In 2003, a National Cancer Institute expert panel concluded that neither spontaneous nor induced abortions are associated with an increased risk of breast cancer (http:// www.cancer.gov/cancerinfo/ere-workshop-report).

2.2.1.4. Breastfeeding

In the 1920s, it was observed that the children of women with breast cancer were less likely to have been breastfed for 1 year than the children of control women (Lane-Claypon, 1926). Lactation decreases risk of breast cancer in parous women, although the overall reduction in risk varies substantially within the population studied

(Lipworth *et al*., 2000). The relative risk for breast cancer decreases by 4.3 percent for every 12 months of breastfeeding (Collaborative Group on Hormonal Factors in Breast Cancer, 2002) and decreased risk with prolonged lactation (Lipworth *et al*., 2000). The risk reduction appears to be greatest among women with high parity, where the risk reduction due to breastfeeding may be as great as 50 percent (Romieu *et al*., 1996) and among premenopausal women with lactation durations ≥2 years, where the breast cancer risk reduction may be 30 percent (Newcomb *et al*., 1994). The magnitude of the decline was consistent across age at breast cancer diagnosis, race or ethnicity, different reproductive patterns, and various personal characteristics. This led to the conclusion that the limited time women in developed countries breastfeed are likely the reasons for high breast cancer incidence (Kelsey *et al*., 1993). Breastfeeding is hypothesized to reduce the risk of breast cancer through differentiation of breast tissue and reduction of the lifetime number of ovulatory cycles. Breastfeeding may result in further terminal differentiation of the breast epithelium, making it more resistant to carcinogenic change. However, some reviews have consistently found no association that breastfeeding reduces risk. There is no solid consensus about the relationship between breastfeeding and breast cancer. Expanded consideration of possible confounders for this relationship is required to determine if breastfeeding is protective and how protection might be conferred (Yang and Jacobsen, 2008).

2.2.2. Family History and Medical History

Hereditability of cancer can be viewed from several angles. Inherited cancer may refer to all cancer cases bearing established causal genetic mutations, and it may refer to cases in families with multiple cancer cases, although no common causal genetic trait has been identified. Since 1860, breast cancer has been recognized to cluster in families as described by the French surgeon Paul Broca (Lynch *et al*., 2008). BRCA1, BRCA2, ATM and TP53 are the most common high penetrance genes exhibiting allelic variants predisposing to hereditary breast cancer (Easton *et al*., 1993). Germline

mutations in BRCA1 and BRCA2 genes are the main genetic and inherited factors for breast cancer. It is estimated that about five to ten percent of all breast cancer cases are caused by mutations in these high-risk genes (Easton, 2002; Lux *et al*., 2006; Lynch *et al*., 2008). These mutations are important in developing early onset and increasing the risk of familial breast cancer, and responsible for 90 percent of hereditary cases (Mahdavi *et al*., 2019). In many families no such pattern can be found, but the history is still indicative of a kind of genetic predisposition. But this susceptibility explains only a small fraction of the familial risk and a much smaller fraction of 5-10 percent of breast cancer cases and decreases markedly with age; approximately 33 percent of cases age 20-29 years compared with approximately 2 percent of cases age 70-79 years (Claus *et al*., 1996). Several studies have been undertaken to find other high penetrance breast cancer susceptibility genes. Genetic variants in CHEK2, PTEN, STK11, CDH1, NBS1, RAD51C, RAD51D, BARD1, PALB2, XRCC2 and XRCC3 have also been implicated in breast cancer risk. Some of these genes are involved in multiple cancer syndromes like Li-Fraumeni (TP53), Peutz-Jeghers (STK11/LKB1) and Cowden syndrome (PTEN) (Ko and Prives, 1996; Nelen *et al*., 1996; Hemminki *et al*., 1998; Sigal and Rotter, 2000; Gasco *et al*., 2003). In contrast to familial cases, low-penetrance genes contribute to sporadic cases of breast cancer that usually appear unilaterally and have a relatively late age at diagnosis (Rebbeck, 1999).

Some studies suggest that hypertension increases the risk of all malignancies (Grossman *et al*., 2001) via hypothesized pathways relating to abnormalities of vascular smooth muscle proliferation, carcinogen binding to DNA, or angiogenesis (Felmeden and Lip, 2001). While some early studies reported a positive association between hypertension and breast cancer (Largent *et al*., 2006), some studies showed no differences even when a follow-up of up to 27 years was done (Peeters *et al*., 2000; Manjer *et al*., 2001; Lindgren *et al*., 2007). A meta-analysis of 30 studies, with totally 11643 cases showed statistically significant association between hypertension and

increased breast cancer risk. A study found increased risk in postmenopausal women than in premenopausal women and Asian population (Han *et al*., 2017). Treatment for hypertension has also been associated with breast cancer risk in several studies although the evidence is inconsistent (Grossman *et al*., 2001). Most observational epidemiologic data do not support an association between antihypertensive use and breast cancer. Furthermore, several large randomized clinical trials showed no association between risk of any cancer and antihypertensive treatment (Lindholm *et al*., 2001; ALLHAT Collaborative Research Group, 2002).

History of diabetes has been associated with increased breast cancer risk (Wu *et al*., 2007). The pathways by which diabetes might cause breast cancer involves the insulin pathway, activation of the insulin-like growth factor pathway, and altered regulation of endogenous sex hormones (Wolf *et al*., 2005). The latter two pathways are thought to be key mechanisms linking obesity and breast cancer. In one study wherein type 1 and 2 diabetes was distinguished, the risk of breast cancer was increased by statistically significant 17 percent among postmenopausal women with type 2 diabetes (Michels *et al*., 2003). Although this is the largest study to date with the longest follow-up (over 22 years), self-reported body weight and lack of data on central obesity could have resulted in residual confounding. In the Cancer Prevention Study II which had a follow-up of 16 years, the relative risk for breast cancer mortality among women with diabetes was a statistically significant with a relative risk of 1.27 (Coughlin *et al*., 2004). A meta-analysis of 40 studies showed there was significant increase in risk associated with diabetes in women. Data from some case control and cohort studies suggest that diabetes carries a moderate increase in the risk of breast cancer (Xue and Michels, 2007).

2.2.3. Lifestyle Factors

Lifestyle factors are modifiable risk factors that can have an important role in primary breast cancer prevention, breast cancer treatment, and tertiary breast cancer

prevention. It is estimated that 90-95 percent of breast cancer cases are connected to environmental factors and lifestyle (Castelló *et al*., 2015). For over three decades, World Cancer Research Fund and American Institute for Cancer Research (WCRF/AICR) has been at the forefront of synthesizing, interpreting, and evaluating the accumulated evidence on the relationship of diet, nutrition, physical activity, and weight with cancer risk (Wiseman, 2008). In the United States alone, approximately 40 percent of all cancer cases could be prevented through health-related choices such as vaccinations and modifiable lifestyle factors, including body weight, physical activity level, alcohol intake, diet, sun exposure, and tobacco use (Islami *et al*., 2018). Some of the major lifestyle factors that are said to increase breast cancer risk are -

2.2.3.1. Anthropometric Factors

Measurements of a human body include height, weight, body mass index (BMI) and other proportions of the body. Even though there is as yet no clear explanation for the connection between height and breast cancer risk, it has been suggested that within populations a 10 cm increase in height corresponds to a 10 percent increase in risk (Hunter and Willett, 1993). Since estrogen plays a key role both in breast cancer development and human growth regulation, the growth spurt has been modestly suggested to influence risk. Estrogen stimulates the pubertal growth spurt and mutations in the ESR1 gene (coding for ERα) have been reported to delay fusion of the epiphyseal plates at puberty (Simm *et al*., 2008; Emons *et al*., 2010). An association between body height and mutations in ESR1 has also been found (Dahlgren *et al*., 2008), and might point towards a more hormone related link.

Body mass index (BMI) and waist-to-hip ratio (WHR) are the most popular measurement for body fat. Several studies have found obesity to be associated with an increased breast cancer risk among postmenopausal women, but in premenopausal women either unrelated or related to a reduced risk (Lubin *et al*., 1985; Willett *et al*., 1985; Chu *et al*., 1991; Pathak and Whittemore, 1992; Ballard-Barbash, 1994;

Franceschi *et al*., 1996; Huang *et al*., 1997; Sonnenschein *et al*., 1999). High BMI, an indicator of obesity, has been suspected to increase risk in postmenopausal women (Hunter and Willett, 1993; Reeves *et al*., 2007). Obesity has been related to both higher endogenous estrogen levels and circulating estrogen in the adipose tissue (Hunter and Willett, 1993; Key *et al*., 2003). In premenopausal women, this connection is unclear (Renehan *et al*., 2008). However, some studies found BMI is inversely associated with risk among premenopausal women (McTiernan, 2003) as obese premenopausal women are more likely to have irregular menstrual cycles and ovulatory infertility (Rich-Edwards *et al*., 1994). Study on association of waist-to-hip ratio with breast cancer risk show significant results (Ballard-Barbash, 1994; Mannisto *et al*., 1996; Hall *et al*., 2000).

2.2.3.2. Physical Activity

Few established breast cancer risk factors are modifiable. However, increasing physical activity and maintaining weight during a woman's adult years offer both individual and population-based opportunities for lowering women's risk of breast cancer (Bernstein *et al*., 1994; De Cree *et al*., 1997). Consistent evidence from epidemiologic studies links physical activity after diagnosis with better breast cancer outcomes as well (Ibrahim and Al-Homaidh, 2011; Chlebowski, 2013). The relationship between physical activity, anthropometric factors, and breast cancer risk may be mediated by several pathways including the steroid hormone, insulin, and insulin-like growth factor pathways. The link between estradiol and breast cancer has been supported by in vitro (McManus and Welsch, 1984; Laidlaw *et al*., 1995) and in vivo (Chang *et al*., 1995) studies showing that estradiol increases the mitotic activity of breast epithelial cells. Physical activity appears to have a direct physiological effect on steroid hormone levels, most clearly during the pubertal and premenopausal stage. Increased physical activity has been directly associated with reduced circulating levels of endogenous estradiol and progesterone among normally cycling women (Shangold

et al., 1979; Ellison and Lager, 1986). It has an indirect effect on exposure to ovarian steroid hormones, in that high level of moderate and vigorous physical activity result in delayed menarche, irregular or anovulatory menstrual cycles, a shortened luteal phase, and in the extreme, secondary amenorrhea (Warren, 1980; Frisch *et al*., 1981; Bernstein *et al*., 1987).

Studies on physical activity and circulating hormone levels in postmenopausal women have not given consistent results (Newcomb *et al*., 1995; Verkasalo *et al*., 2001; Atkinson *et al*., 2004). Breast cancer risk is decreased most with recreational and household activities after menopause. Multiple mechanisms could explain associations between postmenopausal breast cancer, estrogen levels, and physical activity. The first relates to BMI, increase of which is related to breast cancer risk (World Cancer Research Fund and the American Institute for Cancer Research, 2007). This relation might exist in part because after menopause, ovarian estrogen production ceases and adipose tissue becomes a key endogenous source of circulating estrogen (Kendall *et al*., 2007). Hence, by reducing body fat through exercise, estrogen levels may decrease resulting in a lower risk of breast cancer. Levels of adipokines that influence estrogen biosynthesis can also be altered with weight loss (Cleary and Grossmann, 2009).

2.2.3.3. Sleep Cycle and Circadian Disruption

Several studies have investigated a potential link between night shift work and the development of breast cancer (Davis *et al*., 2001; Schernhammer *et al*., 2001, 2006) and have shown an increased risk among women who work in occupations that typically involve some degree of shift work (Megdal *et al*., 2005). The release of nearly all hormones exhibits a circadian timing patterned on approximately a 24-hour cycle. Agents that disrupt circadian rhythm may also alter endocrine function and thereby the regulation of reproductive hormones. Sleep exerts a profound effect on endocrine function and hormones such as melatonin and cortisol (Czeisler and Klerman, 1999). Clinical studies in women with breast cancer showed that they had much less night-

time melatonin levels in urine than a control group of women with no breast cancer (Schernhammer and Hankinson, 2005). It was reported that melatonin is reduced and estrogen elevated in nurses with a history of rotating night shifts (Schernhammer *et al*., 2004). The Finnish Twin Cohort Studies concluded that the risk of breast cancer was lower in women who sleep longer (>9 hr) compared to average sleepers (7-8 hr) (Verkasalo *et al*., 2005).

2.2.3.4. Alcohol

There is substantial evidence to support the association of increased breast cancer risk with alcohol consumption. In a pooled analysis of six cohort studies, alcohol was shown to increase breast cancer risk linearly with alcohol consumption from 1-5 drinks/day (Smith-Warner *et al*., 1998). However, only a modest 15 percent increase in risk was seen in a study of alcoholic women (Kuper *et al*., 2000). A collaborative reanalysis of data from 53 epidemiological studies (Hamajima *et al*., 2002) estimated that the relative risk of breast cancer increased by 7.1 percent for each additional 10 g per day intake of alcohol, i.e., for each extra unit or drink of alcohol consumed daily. The association was seen in both premenopausal and postmenopausal women but does not vary by type of alcoholic beverage (Smith-Warner *et al*., 1998; Tjonneland *et al*., 2007), and does not seem to depend on drinking frequency (Tjonneland *et al*., 2003; Horn-Ross *et al*., 2004). Recent alcohol intake seems to be more relevant than past intake. Alcohol intake in adolescence is not associated with subsequent breast cancer risk (Holmberg *et al*., 1995, Marcus *et al*., 2000). A controlled feeding trials have shown that moderate alcohol intake increases circulating estrogen levels in both premenopausal (Reichman *et al*., 1993) and postmenopausal (Dorgan *et al*., 2001) women.

Recent advances have indicated that alcohol consumption is strongly related to estrogen receptor positive (ER+) breast cancers (Deandrea *et al*., 2008). Alcohol-associated impact on breast cancer appears to be effective in ER+ invasive lobular

carcinoma, but not in ER+ invasive ductal carcinoma (Li *et al*., 2010). Studies suggest that for alcohol drinkers, interactions with GSTM1 and GSTT1 deletion polymorphisms may play an important role in individual susceptibility to breast cancer (Helzlsouser *et al.*, 1998; Park *et al.*, 2000). In vitro, addition of alcohol to breast cancer cells resulted in ER signalling and cell proliferation of ER+ but not ER- cells (Fan *et al*., 2000; Singletary *et al*., 2001).

The mechanism underlying the carcinogenic effect associated with alcohol is not completely understood. However, an increase in the estrogen level in women consuming alcohol has been hypothesized (Reichman *et al*., 1993). Some proposed mechanisms include alcohol induced production of ROS (Wright *et al*., 1999), and increased adduct formation, possibly due to decreased protein expression of detoxification enzymes (Barnes *et al*., 2000). It is also believed that alcohol intake increases mammary tissue exposure to estrogen, induces mutagenesis through its metabolites, increases free radical-mediated DNA damage, and may influence DNA metabolism and gene expression by affecting one-carbon metabolism (Dumitrescu and Shields, 2005; Seitz and Stickel, 2007).

2.2.3.5. Tobacco

Tobacco is known to contain a variety of compounds that are carcinogenic. Cigarette smoking is the most common form of tobacco use worldwide. Tobacco consumption has been clearly implicated in the causation of many cancer types. However, studies on breast cancer risk have reported positive, inverse and null associations. Despite mixed result, there is growing evidence that smoking may slightly increase the risk of breast cancer. In a meta-analysis, current and former smoking were weakly associated with breast cancer risk; a stronger association was observed in women who initiated smoking before first birth (Gaudet *et al*., 2013). Statistically significant effects have been seen for early age at starting, and for heavy, current, and passive smoking. In some studies, there is significant increase in breast cancer risk

in never smokers with longterm exposure to passive smoking, while other scientist rejects the evidence of this association as inconsistent (Ambrosone and Shields, 1999; Wartenberg *et al*., 2000). In a study where non-smoking women exposed to passive or secondhand smoke were excluded, there was evidence of positive associations between breast cancer and cigarette smoking (Morabia, 2002). In addition, studies suggest that the risk of breast cancer associated with smoking might be increased for premenopausal women (Khuder *et al*., 2001) or women who started smoking in their mid-teens or earlier, or before their first full-term pregnancy (Khuder *et al*., 2001; Terry and Rohan, 2002; Cui *et al*., 2006; Ha *et al*., 2007). Similarly, women who inherited specific variants in genes involved in the metabolism of carcinogens found in tobacco might experience higher risks associated with smoking cigarettes (Morabia, 2002; Terry and Rohan, 2002; Slattery *et al*., 2008).

Tobacco smoke consists of more than 7000 chemical compounds, and more than 60 of these are known carcinogens (Das, 2003). These are transported by the blood stream, deposited and metabolically activated in the breast and surrounding adipose tissues to the breast, which can further be detected in the nipple discharge or as smoking specific DNA adducts in breast tissue (Terry and Rohan, 2002). The most important carcinogens in tobacco smoke are polycyclic aromatic hydrocarbons (PAHs), aryl amines, heterocyclic aromatic amines (HAAs), and N-nitrosamines (Ambrosone and Shields, 1999; Bartsch *et al*., 2000). The ingested or inhaled PAHs are converted to water soluble derivatives mainly via oxidative activation by cytochrome P450 1A1 (CYP1A1) followed by detoxification by phase II enzymes such as glutathione S-transferases (GSTs). PAHs have been shown to be mutagenic to breast cell lines, and as lipophilic compounds they are stored in adipose tissues (Li *et al*., 1996).

Smokeless tobacco products contain more than 3000 chemicals and 28 numerous carcinogens. DNA binding and mutations are among the mechanisms clearly implicated in carcinogenesis due to smokeless tobacco use. Smokeless tobacco generally comes

in the form of chewing tobacco. Smokeless tobacco also generates reactive oxygen species, oxidative stress,and associated DNA fragmentation in laboratory experiments. The major and most abundant group of carcinogens are the non-volatile alkaloid-derived tobacco-specific N-nitrosamines (TSNA) and N-nitrosoamino acids. Other carcinogens reportedly present in smokeless tobacco include volatile N-nitrosamines, certain volatile aldehydes, traces of some polynuclear aromatic hydrocarbons such as benzo[a]pyrene, certain lactones, urethane, metals, polonium-210 and uranium-235 and -238 (Brunnemann and Hoffmann, 1992). Compared to smoking, higher levels of nicotine can enter systemic circulation from smokeless tobacco which indicate a much more potent effect through this route (Li *et al*., 2018). *Tuibur*, a unique form of aqueous extract of tobacco exhibited significant toxicity by reducing the root growth of Allium bulbs and inducing tumor formation in the roots. This form of smokeless tobacco is most commonly used in the north eastern part of India like Manipur, Mizoram, Sikkim and Tripura. It is used for gargling, or sipped and spitted out only after it becomes diluted (Mahanta *et al*., 1998).

2.2.3.6. Betel Quid

Betel quid can contain a variety of ingredients and combinations depending on different parts of the world. Usually, it contains a mixture of areca nut (*Areca catechu*), catechu (*Acacia catechu*) and slaked lime (calcium oxide and calcium hydroxide) and several condiments according to taste preference, wrapped in betel leaf (*Piper betle*), some add tobacco (Nair *et al*., 2004). In Mizoram, the most common preparation used in betel quid is a mixture of areca nut wrapped in betel leaf with a paste of slaked lime. In vitro and in vivo experiments have shown that betel quid consumption can cause micronuclei and DNA adducts formation, chromosomal aberrations, allelic imbalances and sister chromatid exchange in oral mucosa cells (IARC, 2004). Calcium hydroxide a major content of slaked lime in the presence of areca nut is responsible for the formation of ROS (reactive oxygen species) known to cause oxidative damage

in the DNA of buccal mucosa cells of betel quid chewers (Nair *et al*., 2004). A study in Assam, India, reveals that betel quid chewers have higher risk of having breast cancer than the non-chewers (Rajbongshi *et al*., 2015). In a case control study among the Mizo population betel quid chewing was found to be a significant risk factor for developing breast cancer. Multifactor dimensionality reduction identified betel quid chewing as the single main risk factor and women with betel quid chewing history have five times the risk of developing breast cancer (Kaushala *et al*., 2010). Higher total number of genomic alterations were seen in breast cancer tumors of betel quid than to non-chewers (Kaushal *et al*., 2012).

2.2.4. Diet and Nutrition

The human diet contains a great variety of natural carcinogens and anticarcinogens (Sugimura, 2000). Many of these may act through the generation of oxygen radicals, which in turn may lead to DNA damage. There is an almost universal agreement that diet or nutritional practices in some form must play a role in establishing breast cancer risk. However, no specific component of the adult diet and no particular nutrient have been consistently associated with breast cancer risk (World Cancer Research Fund and American Institute for Cancer Research, 2010). The results of a large meta-analysis of 26 published studies from 1982 to 1997 (Gandini *et al*., 2000) and of a pooled analysis of 8 cohort studies (Smith-Warner *et al*., 2001b) suggest that fruit and vegetable consumption during adulthood is not significantly associated with reduced breast cancer risk. A pooled analysis of individual data from seven prospective studies in four countries comprising 337,819 women and 4980 breast cancers also suggested a lack of association between total fat, saturated fat, mono and poly unsaturated fat intake and breast cancer risk (Hunter *et al*., 1996).

However, some studies have shown that a high intake of fat, especially that of poly unsaturated fatty acids, has been shown to increase breast cancer risk (Bartsch *et al*., 1999), while intake of fruits and vegetables, sources of natural antioxidants, has

been shown to decrease the risk (Lee, 1999; McKeown, 1999). Consumption of meat has been associated with increased breast cancer risk in some, but not all studies (Zheng *et al*., 1998). Dietary fat has long been suspected to be the reason for this association, but recent studies support the role of heterocyclic aromatic amine (HAAs) found in well-done meat. HAAs require metabolic activation by N-acetyltransferase to be able to exert their harmful effects (Hein *et al*., 2000). Soy, or more specifically genistein, with a chemical structure like steroidal estrogens, has been shown to have both anti-carcinogenic and cancer promoting effects (Bouker and Hilakivi-Clarke, 2000).

The Mizos are a unique group of people differing from mainland India in their culture, lifestyle and dietary habits. Some of the indigenous foods of the Mizos contained ash filtrate (*ching-al*), smoked or sun-dried or fermented meats and vegetables. Studies on the association of stomach cancer risk with these traditional food reveals that frequent consumption of fermented pork fat (*sa-um*), smoked dried salted meat and fish elevates risk. The use of soda or ash filtrate, used as a food additive, increased the risk of stomach cancer (Phukan *et al*., 2006). Studies reveals that smoked meat, smoked fish and soda also increased lung cancer risk (Phukan *et al*., 2014). Even though there is no published data on association of breast cancer risk among the Mizos with fermented soyabean (*bekang*), a study in a Japanese population-based cohort showed no association with risk (Shirabe *et al*., 2021).

2.2.5. Exogenous Hormones - Oral Contraceptives and Postmenopausal Hormone

Sex hormones have become one of the most widely used drugs among women. It is not possible to study the effects of single hormone as many are used in the same patient either in combination or consecutively. Therefore, risks are generally assessed in relation to the therapeutic goal of the treatment, i.e., oral contraception or hormone replacement therapy (HRT). Results from studies on the role of oral contraceptives to breast cancer proneness have been somewhat conflicting. However, data from 54

studies concluded that current use of oral contraceptives poses a slight (24%) increase in the risk, which disappears 10 years after the cessation of use (Collaborative Group on Hormonal Factors in Breast Cancer, 1996). The importance of progesterone in breast cancer risk has been highlighted by several recent observational studies and most notably the Women's Health Initiative randomized trials that have shown that combined estrogen plus progestin hormone therapy increases breast cancer risk while use of estrogen alone does not (Ross *et al*., 2000; Chlebowski *et al*., 2003).

Results from 51 studies indicated that risk of having breast cancer is slightly increased in women using HRT (Collaborative Group on Hormonal Factors in Breast Cancer, 1997). A 35 percent increase in risk was seen for women who have used HRT for 5 years or longer, being comparable with the effect of delaying menopause, and the increase largely disappeared 5 years after terminating the use of hormones. The combined estrogen-progestin regimen is associated with greater increase in risk than estrogen alone (Schairer *et al*., 2000). Moreover, in some studies, cancers in women who have ever used HRT tend to be less advanced at diagnosis and biologically less aggressive than those in never users (Holli *et al*., 1997), but contrasting results exist (Stallard *et al*., 2000). The overall mortality among HRT users has been shown to be lower but the benefit diminishes with longer duration of use (Grodstein *et al*., 1997). Xenoestrogens include pesticides, dyes, pollutants, plasticizers and food preservatives that have estrogen-like effects, and they have been shown to have a role in the etiology of breast cancer (Davis *et al*., 1993; Safe, 1997; Garner *et al*., 2000a; 2000b; Spink et al., 2000). Xenoestrogens have also been called endocrine disrupters because they interfere with the actions of endogenous estrogen. For instance, catechol metabolites of polychlorinated biphenyls (PCBs) have been suggested to alter estrogen metabolism by inhibiting the inactivation of carcinogenic estrogen metabolites (Garner *et al*., 2000a).

2.2.6. Environmental and Occupational Exposures

It has been widely suggested that environment may play an important role in increasing breast cancer incidence. Environmental pollutants such as organochlorines, polycyclic aromatic hydrocarbons, dioxins, and bisphenol A and extremely low frequency (ELF) magnetic fields have been linked with breast cancer risk in animal studies and may plausibly be associated with risk in humans. Exposure to some naturally occurring trace elements and heavy metals are suspected to influence breast cancer risk.

2.2.6.1. Environmental Pollutants

The International Agency for Research on Cancer (IARC) considers several environmentally abundant chemicals, chemical compounds, and their metabolites to be either known (IARC, 1997b) or suspected (IARC, 1997a) human carcinogens. Among these, organochlorines, polycyclic aromatic hydrocarbons (PAHs), dioxins, and bisphenol A (BPA) have received particular attention with respect to breast cancer. The specific mechanisms by which exposure to environmental pollutants could impact breast cancer risk are varied but their persistence in the environment and their tendency to accumulate in adipose tissue, including the fatty tissue in the breast is common. Concerns that exposure to these pollutants could influence risk stem primarily from the fact that many of these chemicals are 'endocrine disruptors,' mimicking or blocking the effects of specific hormones (Rudel *et al*., 2007). Since some of these pollutants mimic the activity of estrogen, it is hypothesized that they could influence the initiation or progression of breast cancer in humans through estrogenic effects (Soto *et al*., 1995; Connor *et al*., 1997; Shekhar *et al*., 1997).

Dichloro-diphenyl-trichloroethane (DDT) and its associated compounds have been classified as Group 2B carcinogens (possibly carcinogenic) by IARC (IARC, 1997a). It is ubiquitous in nature and accumulates in the food chain, particularly in

fish and fatty foods and has estrogenic effects (Soto *et al*., 1995; Shekhar *et al*., 1997). Polycyclic aromatic hydrocarbons (PAHs) are formed as by-products of combustion from tobacco smoke, air pollution, vehicle exhaust, and smoked or grilled meat and fish (Brody *et al*., 2007). Some PAHs are considered environmental estrogen, although their estrogenic effects are generally weak (Santodonato, 1997).

Although some studies provide no strong evidence of gene-environment interactions, there is some evidence to suggest that women with a GSTM1 null phenotype (Rundle *et al*., 2000) may be susceptible to the effects of PAH exposure. Studies have suggested that exposure to high levels of PAHs in early childhood (Bonner *et al*., 2005), at the time of first birth (Nie *et al*., 2007), or in the past 10-20 years (Lewis-Michl *et al*., 1996) could increase risk of postmenopausal breast cancer; associations with risk of premenopausal breast cancer are somewhat less consistent (Bonner *et al*., 2005; Nie *et al*., 2007). Bisphenol A (BPA) is an important monomer in the production of the epoxy resins that line food and beverage cans and in the production of the shatter proof polycarbonate plastics that are used in a wide variety of household products and devices. Studies in mouse and rat models have suggested that in utero exposure to BPA results in alterations in the architecture of the adolescent and adult breast. Mammary glands in animals prenatally treated with BPA have an increased number of undifferentiated epithelial structures, more progesterone receptor-positive (PR+) epithelial cells, decreased apoptosis and enhanced sensitivity to estradiol (Durando *et al*., 2007; Murray *et al*., 2007; Moral *et al*., 2008).

2.2.6.2. Radiation – Ionizing and Non-Ionizing

It is clearly established that exposure to ionizing radiation is an important risk factor for breast cancer. The most important information available regarding association of ionizing radiation with breast cancer risk comes from studies of the survivors of the atomic bombings of Hiroshima and Nagasaki (Land *et al*., 2003) and long-term follow-up of cohorts of people receiving radiation exposure for medical

reasons (Horwich and Swerdlow, 2004). Age at exposure appears to be an important risk determinant and exposure around the time of puberty conferring the highest risk (Goss and Sierra, 1998). Ionising radiation has been shown to increase breast cancer risk among female flight attendants, nurses, chemists and insulators (John and Kelsey, 1993; McCormick, 1999; Weiderpass *et al*., 1999). Even though there is no strong evidence, electromagnetic fields have also been hypothesised to affect breast cancer risk by suppressing melatonin production (Kheifets and Matkin, 1999). Other occupational studies among pharmaceutical industry workers, cosmetologists, beauticians, chemists, teachers, social workers, and cashiers have found association with breast cancer risk (Goldberg and Labreche, 1996; Welp *et al*., 1998).

McDowall defined electromagnetic field exposure as living within 30 m of either electrical installation equipment or an overhead power cable (McDowall, 1986). Experimental studies in laboratory animals supports a link between Extremely low-frequency (ELF) magnetic field exposure and decreased melatonin levels but limited data support this link in humans (Davis *et al*., 2006). It has been hypothesized that disruption of the normal nocturnal rise in melatonin resulting from exposure to ELF magnetic fields could increase breast cancer risk (Stevens *et al*., 1992). However, a large number of studies found no evidence that exposure to residential magnetic fields is associated with increased risk (Davis *et al*., 2002; London *et al*., 2003). A meta-analysis of 15 case control studies from 2000 to 2009, involving 24,338 cases and 60,628 controls, found no significant association between breast cancer risk in relation to ELF-EMF exposure, even when stratifying by menopausal status or the source of exposure (Chen *et al*., 2010).

2.2.6.3. Trace Elements and Heavy Metals

Trace elements and heavy metals occurring naturally in the environment may influence a woman's risk of developing breast cancer. Exposure to naturally occurring trace elements and heavy metals can be from a variety of sources with geographic

variation, including drinking water, air, food, and occupational exposure. Some trace elements such as arsenic (IARC, 1987) and some heavy metals such as cadmium (IARC, 1993) and lead (IARC, 1987) are considered by IARC to be either known or suspected human carcinogens at specified doses of exposure. Others, such as selenium, copper, iron, and zinc, may plausibly be associated with breast cancer risk given their biological roles. However, evidence associating exposure to these elements with breast cancer risk is limited (Navarro Silvera and Rohan, 2007). Association studies with exposures to heavy metals is also limited and predominantly null, but some studies have found positive association between breast cancer risk and exposure to cadmium (McElroy *et al*., 2006). However, studies on the carcinogenicity of selenium and arsenic does not provide clear evidence to support that they increase breast cancer risk (Garland *et al*., 1996; Navarro Silvera and Rohan, 2007). Copper and iron are biologically important in the production of reactive oxygen species, but excessive exposure to these metals could contribute to oxidative stress and, potentially carcinogenesis. But studies do not give proper association between breast cancer risk and these two metals.

2.3. Mitochondrial DNA

The mitochondrial genome though small is responsible for ensuring that the powerhouses of our cells function properly. As a by-product of their role as powerhouses of our cells, mitochondria generate reactive oxygen species (ROS). This circular genome is more in quantity than its nuclear counterpart and has a higher mutation rate than the nuclear genome and represents less than 1 percent of total cellular DNA. Mammalian mtDNA contains no introns and lacks histones. This, along with its close proximity to ROS produced through oxidative phosphorylation in the mitochondria, make mtDNA vulnerable to oxidative damage and mutations. In fact, the mutation frequency in mtDNA is approximately tenfold greater than that in nuclear DNA (Johns, 1995; Grossman and Shoubridge, 1996). Mutations in the mitochondrial genome have been associated with diverse forms of human disease. Over the years, a

probable role for mitochondria in both aging and cancer has emerged. ROS production has been proposed to cause somatic mitochondrial mutations. A large body of evidence suggests that somatic mtDNA mutations play a role in breast carcinogenesis.

2.3.1. Displacement Loop

The d-loop region has been shown to be a mutational 'hot spot' in human cancer. The incidence of somatic mutations in the d-loop region is found in all tumors examined to date and appears to be a universal feature of all cancers (Modica-Napolitano *et al*., 2007). Studies show somatic mutations in majority of breast cancer patients and most of the mutations identified were in the d-loop region (Parrella *et al*., 2001; Tan *et al*., 2002). In a cohort study focusing on the d-loop, 36.36 percent of samples presented somatic mutations while 90.91 percent of samples showed germline mutations (Barekati *et al*., 2010). Within this region, a poly-C repeat stretch, named D310, contained the majority of mutations (Santos *et al*., 2012; Xu *et al*., 2012). D310 alterations were more frequent in cervical cancer followed by bladder cancer, breast cancer and endometrial cancer (Parrella *et al*., 2003). Among the Chinese, the germline polymorphism of T16189C is suggested to convey increased risk considering the high frequency observed in breast cancer patients (Wang *et al*., 2006). A study among non-Jewish European American found variants 12308G and 10398G to increase breast cancer risk (Covarrubias *et al*., 2008). In another study d-loop mutations were associated with advanced age (>50 years), negative estrogen and progesterone receptor status, as well as poorer disease-free survival (Tseng *et al*., 2006). A study from China of D310 mutations in familial breast cancer recorded extremely high frequencies (Yu *et al*., 2008). The first study on breast cancer and association with mtDNA among the Mizos found mitochondrial gene alterations may attribute for risk (Ghatak *et al*., 2014).

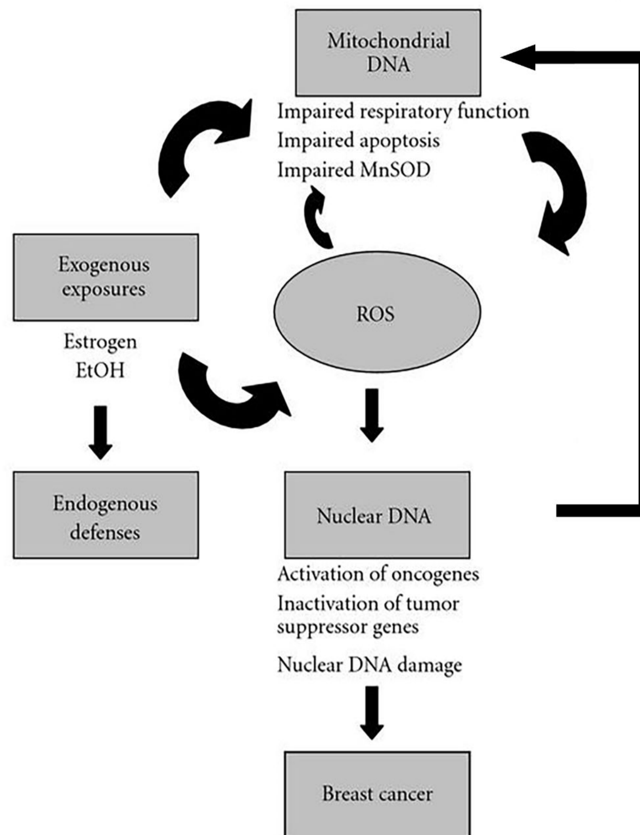


Figure 1. Schema showing how ROS may affect mitochondrial and nuclear DNA leading to breast carcinogenesis (Rohan et al., 2010).

2.3.2. Cytochrome C Oxidase

Cytochrome c oxidase is a large integral membrane protein which is encoded in the mitochondrial genome. It is a terminal oxidase of the mitochondrial electron transport chain, and is expressed in the mitochondrial inner membrane. It is responsible for production of a critical enzyme that controls mitochondrial respiration and is central to apoptosis (Payne *et al*., 2005). This enzyme is the terminal electron acceptor in the electron-transport chain and catalyses the complete reduction of molecular oxygen to water with the supply of four electrons from cytochrome c and four protons taken up from the mitochondrial matrix. Reactive oxygen species (ROS) are known for transient existence, induce oxidative damage leading to both nuclear DNA and mtDNA aberrations, and thus play an important role in carcinogenesis. Increased ROS

generation may alter signal transduction pathways, resulting in activation of oncogenes or inactivation of tumor suppressor genes. Defects in the mitochondrial genome and function are suspected to contribute to the development and progression of cancer (Ye *et al*., 2008). A depletion of COI subunit has been observed in breast cancer (Putignani *et al*., 2008). Several cancer associated mutations found in mtDNA result in structural modifications of cytochrome c oxidase (Namslauer and Brzezinksi, 2009).

2.4. Glutathione S-transferases

In 1961, an enzymatic reaction responsible for the first step in the conjugation of xenobiotics with glutathione was recognized (Combes and Stakelum, 1961). Glutathione s-transferases (GSTs) represent an important group of enzymes encoded by a superfamily of GST genes that detoxify both endogenous compounds and foreign chemicals such as pharmaceuticals and environmental pollutants. Allelic variations are found in genes encoding for these GSTs (Board *et al.*, 1997; Blackburn *et al*., 2000; Strange *et al*., 2000). In estrogen metabolism, GSTs play a role in the catalysis of glutathione (GSH) conjugation of catechol estrogen quinones, the reactive intermediates of estrogen metabolism capable of binding to DNA (Raftogianis *et al*., 2000). GSTs may also be involved in the activation of some carcinogens such as halogenated hydrocarbons (Taningher *et al*., 1999; Strange *et al*., 2000). Human tissues show differential expression of the multiple forms of GSTs (Johansson and Mannervik, 2001). The absence of specific isoenzymes affects the tolerance of organisms to chemical challenges and may result in an increased rate of somatic mutations and higher susceptibility to disease. The ability of many tumours to exhibit increased levels of intracellular GST expression has been linked to mechanisms of chemotherapeutic drug resistance (Black *et al*., 1990). Carriers of homozygous deletions in the GSTM1 and GSTTI genes have an absence of GST mu and GST theta enzyme activity, respectively (Seidegard *et al*., 1988, Hallier *et al*., 1993, Pemble *et al*., 1994). These deletion variants have been useful for molecular epidemiological

studies of cancer because they divide study subjects into two well-defined susceptibility classes: those who are and those who are not able to detoxify potential carcinogens by the metabolic pathways regulated by GSTM1 and GSTTI. Studies among the Mizos indicate GSTM1 and GSTT1 null genotypes were associated with an increased risk of breast cancer (Kimi *et al*., 2016).

2.4.1. Glutathione S-transferase Mu (GSTM1)

GSTM1 is located in the middle of a cluster of five mu class genes on chromosome 1p13.3 (Pearson *et al*., 1993). The homozygous deletion (null genotype) of the GSTM1 gene leads to the total absence of the respective enzyme activity (Seidegard *et al*., 1988). The frequency of the null genotype is around 50 percent in Caucasians and Asians, but only 27 percent in Africans (Garte *et al*., 2001). Allelic variants of GSTM1, A and B have also been found to exist, but based on current knowledge they have no consequences in the catalytic activity of the enzyme (Widersten *et al*., 1991). In addition, gene duplication has been found to exist in Saudi Arabians (McLellan *et al*., 1997). The GSTM1 genotype has been examined in relation to individual breast cancer risk in several studies. Some of these studies pointed to an association between GSTM1 null genotype and breast cancer risk in postmenopausal women (Helzlsouer *et al*., 1998; Charrier *et al*., 1999). But a large number of studies conducted did not find any link between GSTM1 null genotype with breast cancer risk (Kelsey *et al*., 1997; Bailey *et al*., 1998; Ambrosone *et al*., 1999; Millikan *et al*., 2000). An increased risk for premenopausal women has also been shown, but only in one study (Park *et al*., 2000). Despite these discrepant findings, one meta-analysis suggested that the GSTM1 null genotype poses a moderately increased risk for postmenopausal breast cancer (Dunning *et al*., 1999). Furthermore, the risk has been shown to be modified by BMI (Helzlsouer *et al*., 1998), family history (Millikan *et al*., 2000), use of alcohol (Park *et al*., 2000), and smoking (Millikan *et al*., 2000). The GSTM1 null genotype was not found to be associated with tumour characteristics or survival in one study (Lizard-

Nacol *et al*., 1999) but has been suggested to be associated with both longer (Kelsey *et al*., 1997) and shorter (Nedelcheva Kristensen *et al*., 1998) survival. In the study showing a reduced survival time, the concurrent presence of both the GSTM1 null and GSTT1 null genotypes was associated with positive lymph node status. A strong association has also been shown between GSTM1 deletion and increased PAH-DNA adduct levels in breast tumour tissue (Rundle *et al*., 2000).

2.4.2. Glutathione S-transferase Theta (GSTT1)

The human GSTT1 gene is localised on chromosome 22q with the GSTT2 gene and a pseudogene (Pemble *et al*., 1994; Tan *et al*., 1995). Like GSTM1, homozygous deletion of the GSTT1 gene, leading to total absence of the respective enzyme activity has been observed. Large inter-ethnic differences have been reported in the frequencies of the GSTT1 null genotype being significantly lower among Caucasians (20%) compared to Asians (60%) (Nelson *et al*., 1995). GSTT1 is expressed in human erythrocytes, and various tissues including liver but no expression in breast tissue has been reported (Pemble *et al*., 1994; Landi, 2000). In most of the studies on GSTT1 genotype and breast cancer risk, no significant association has been found (Bailey *et al*., 1998; Helzlsouer *et al*., 1998; Curran *et al*., 2000; Millikan *et al*., 2000; Park *et al*., 2000). However, in one study the risk was found to be modified by the use of alcohol (Helzlsouer *et al*., 1998), and in another study a remarkably lower risk was suggested for premenopausal women lacking the GSTT1 gene (Garcia-Closas *et al*., 1999).

There are only a few published data on breast cancer among the Mizo population. One study reported breast cancer to be cause by betel quid use (Kaushala et al., 2010), and another on GSTM1 and GSTT1 null genotypes (Kimi et al., 2016). It was also reported that mitochondrial gene alterations may attribute breast cancer risk (Ghatak et al., 2014). Some of the indigenous foods of the Mizos contain ash filtrate (*ching-al*), smoked or sun-dried or fermented meats and vegetables. Studies on the association of stomach cancer risk with some of the traditional food reveals elevated risk (Phukan et

al., 2006, 2014); but there are no data available to associate or even to dissociate. Like a study in Japan showed no association of fermented soyabean (bekang – a delicacies of the Mizos) with risk (Shirabe et al., 2021). With breast cancer incidence on the rise and with only a few data available, it is of great importance to identify the specific risk factors prevalent in this region. The Mizos with their unique lifestyle and dietary habits, the reports from other parts of the world are sometimes not applicable or sufficient, or sometimes not acceptable. Since breast cancer etiology is multifactorial, and differs geographically and ethnically, the question remain what could be the rise in incidence. Is there any change that was not there 50 years or so before. With this in mind, this study was designed to include genes related to catabolism and detoxification of xenobiotics, any mitochondrial gene alterations and any regional characteristics that could influence risk; or whether risk increase when there is association of gene alterations with any risk factors.

3. Objectives

i) To determine the potential demographic risk factors associated with breast cancer in Mizo population.

ii) Study of mitochondrial control region (D-loop), cytochrome c oxidase subunit I (CO1) sequence variations associated with breast cancer risk.

iii) Genes related to catabolism and detoxification of xenobiotics (GSTM1 and GSTT1) to explore their contribution for breast cancer.

4. Materials and Methods

4.1. Data Source and Sample Collection

A retrospective case-control study was conducted with the approval of the Institutional Ethics Committee, Civil Hospital Aizawl [No. B 12018/1/13-CH(A)/IEC/28 of October 15, 2014]. The study was initiated in 2014 with a follow-up after 5 years in 2020. Case data was collected from breast cancer patients registered between 1998-2014, a period spanning for 17 years. Control data was collected from outdoor patient who visited Aizawl Civil Hospital and from personal invitation. Informed consent was obtained before each interview as instructed by the Institutional Ethics Committee.

Data was collected from medical records corroborated with a structured questionnaire. Criteria for inclusion in the case study includes confirmed breast cancer registered at Mizoram State Cancer Institute, Zemabawk and Population Based Cancer Registry, Aizawl Civil Hospital; Mizo female with proper contact information for follow-up. All the 758 registered cases were not included in this study because of not meeting the criteria mentioned above, only 363 cases were included. The normal controls were volunteers from outdoor patient who visited Aizawl Civil Hospital and from personal invitation. The criteria for inclusion in the control group was that the volunteer should be a Mizo female at least 18 years of age or above, free of cancer and with no history of cancer in the family, 405 controls matching the criteria mentioned above were included in the study.

4.2. Questionnaire

The questionnaire was structured in such a manner that established as well as suspected risk factors were included. The questionnaire as mentioned in Table 1 included demography, lifestyle, environmental exposure, medical history, reproductive history, family history and anthropometric factors. The questionnaire was specifically

formed to study education, employment, sleep pattern and duration, physical activity, age at menarche, age at first childbirth, total number of children, duration of breast feeding (in months), height (feet), and weight (kilogram) at the time of breast cancer diagnosis or for controls at the time of interview. Both cases and controls were subjected to the same questionnaire format except for questions about breast cancer. Residence of 5 years in an area before having breast cancer for cases was taken as residence even if they had moved to other areas during interview. The interview was conducted in the local language and took about 30-40 minutes for each volunteer. Open-ended question format was used and later categorized. This kind of format does not require a response based on a specific list of choices and allows the participants to answer freely without fear of being wrong. To facilitate comparison, variables were later categorized in two or three or four. For habits, 'quit' is when they have stopped the habit for at least 5 years, for cases before diagnosis and control before interview. Volunteers were recorded as being passive smokers only when there was constant known exposure to secondhand smoke either at work place or at home.

The questions relating to dietary composition and nutrition had a short coming as the subjects were susceptible to both recall bias and selection bias especially against fondness of the food. As such, the questionnaire was structured on frequency of consumption rather than amount which could still be influenced by selection bias especially against fondness of the food. But they do provide a rough idea of their daily dietary intake, but not a prolonged accurate dietary habit. Less is when consumption frequency is trice or less in a month; moderate when they consume once to four times in a week; high is when they consume it at least five times or more in a week.

Total time of breastfeeding was calculated as mean time of breastfeeding multiplied with parity. To calculate mean duration total time of breastfeeding was divided by number of children. Job and age gap between children was also considered but 15 or 18 months given if response is more than a year, unless specific mention of

Table 1. Interview questionnaires.

Category	**Variables**
Personal information	Name, age, residence, occupation, education, handedness, weight, height
Lifestyle	Behavioral habits (including secondhand smoke), diet, sleep, physical activity
Environmental exposure	Electromagnetic and pesticide exposures
Medical history	All major illness including diabetes mellitus, hypertension, any cancer
Menstrual history	Menarche and menopause ages, dysmenorrhea, menopausal problems
Reproductive history	Marriage, parity, breastfeeding history, use of contraceptives
Family history	Breast cancer history (1st degree, 2nd degree and 3rd degree), maternal age
Knowledge	Self-examination, clinical examination, breast cancer-related knowledge

time (Butt *et al*., 2014). All live births were included in the study as death of a child after birth had to be considered in some of the volunteers.

4.3. Anthropometric and Physical Activity Measurement

Body Surface Area was calculated using Mosteller's formula which is the square root of weight (kilogram) times the height (centimeter) divided by 3600. Body Mass Index was measured as weight in kilograms (kg) divided by the square of height in meters (m^2). Although an imperfect measure, BMI is highly correlated with percentage of body fat (Deurenberg *et al*., 1991). The World Health Organization has defined the following cut points for BMI: BMI less than 18.50 is considered underweight; BMI between 18.50 and 24.99 is described as normal or healthy; BMI between 25.00 and 29.99 is grade 1 overweight or overweight; BMI between 30.00 and 39.99 is grade 2 overweight or obese; BMI greater than or equal to 40.00 is grade 3 overweight or morbidly obese. For our study, quartile distributions were used to categorize BMI to facilitate comparison; underweight = BMI less than 18.50; healthy = 18.50-24.99; overweight = 25.00-29.99; obese = 30.00-39.99 (WHO, 1995).

To quantify physical activity, metabolic equivalent of energy expenditure (MET) value was assigned to each reported activity according to the Framingham Physical Activity Index score (Dorgan *et al*., 1994). MET was calculated by summing the number of hours spent in each activity intensity level and multiplying by a respective weight factor derived from the estimated oxygen consumption requirement for each intensity level. One MET being equivalent to the amount of energy a person expends at rest. In this study, MET value was measured for a typical 24-hour duration, categorizing by intensity with slight modification. Since sleep duration was recorded during interview, score of 1 MET was multiplied with the hours spent sleeping or at rest as shown in Table 2. Depending on their main occupation, age and health condition, 5 or 7 hours was multiplied for 'sedentary' for all subjects, the rest of the 24 hours was divided between their main occupation, recreational and household activities. Even though farming is categorized under 'heavy', the number of hours spend for the category differs depending on their age and health condition. If the subject was healthy and under 50 years of age, 5 hours was multiplied. Even if subject was under 50, 2 hours multiplied under 'heavy' if they were occasional farmers. If aged between 51 and 65 years, 2 hours was multiplied, and 1 hour if between age 66 and 69 years, and none if 70 years or more. Physical activity was dichotomized into normal and heavy, where light and moderate MET was grouped as normal; and heavy as heavy to enable further analysis.

Table 2. Physical activity of 24-hour duration.

Sleep (1 x hours)	Sedentary (1.1 x hours)	Slight (1.5 x hours)	Moderate (2.4 x hours)	Heavy (5 x hours)
	Personal time	Knitting	Household work	Farming
	Meal time	Desk job/Teaching	Carpentry	Sports
	Watching television	Shopkeeping	Nursing/beautician	
	Riding in a car	Tailoring	Walking/ Gardening	
	Quite time/meditation	Handloom weavers	Pig/poultry/cattle farm	

4.4. Molecular Study

For molecular study, 49 cases and 41 controls were selected. Healthy controls were individually matched with cases by age (cases = 48.96±11.33, control = 46.73±10.21) except for a 78-year-old female case. Matching area status of incidence, (cases urban=26, rural=23; control urban=25, rural=16) could not be done in spite of inviting residents of Ramlaitui, a rural area of Lunglei District. But to at least match standard of living like farming, residents of Tlangnuam, Aizawl District were personally invited to volunteer for the study (Physical activity score of cases = 36.22±3.12, control = 36.17±2.41; BMI of cases = 23.61±3.65, control = 23.12±3.16).

4.4.1. DNA Isolation

Signed informed consent was obtained from all subjects prior to collection of blood samples. Peripheral blood sample was collected in a 2 ml EDTA vacuum tube and stored at -20ºC. DNA isolation from blood was done using a commercially available kit from Qiagen (Qiagen DNeasy Blood & Tissue Kit, Cat. No.69504). Isolation was done as per instructions given along with the kit. Into a 1.75 ml microcentrifuge tube, 20 µl of proteinase K and 150 µl of anticoagulated blood was added. The volume was adjusted to 200 µl with PBS (pH 7.4). The sample vial was inverted before pipetting out blood. Buffer AL (without added ethanol) of 200 µl was added and mixed thoroughly by vortexing for 10-20 seconds, centrifuged at 8000 rpm for 1 minute so that no solution sticks on the cap. After the sample vial was incubated at 56°C for 10 minutes, 200 µl ethanol (96–100%) was added, and mixed thoroughly by vortex and centrifuge. Lysate was transferred in DNeasy mini spin column placed in a 2 ml collection tube without wetting the rim. Centrifuged at 8000 rpm for 1 min. Flow-through and collection tube was discarded. The DNeasy mini spin column was placed in a new 2 ml collection tube and 500 µl Buffer AW1 was added. Centrifuged for 1 min at 8000 rpm, flow-through and collection tube discarded. The DNeasy mini spin column was placed in a new 2 ml collection tube and 500 µl Buffer AW2 was added.

Centrifuged for 3 min at 14,000 rpm to dry the DNeasy membrane. Flow-through and collection tube discarded. The DNeasy mini spin column was placed in a clean 1.5 ml microcentrifuge tube. Buffer AE of 200 µl was added directly onto the DNeasy membrane. Incubated at room temperature for 1 min, and then centrifuged for 1 min at 8000 rpm to elute. The extracted DNA was stored in at -20ºC until further use.

4.4.2. Amplification of Gene

Table 3. List of primers and sequences.

Gene	Primer Sequence 5′ – 3′
D-loop	FW: AAGACTCGGCAGCATCTCCACACCATTAGCACCCAAAGCT RW: GCGATCGTCACTGTTCTCCACTGTTAAAAGTGCATACCGCCA
CO1	FW: AAGACTCGGCAGCATCTCCATCAACAAATCATAAAGATATTGG RW: GCGATCGTCACTGTTCTCCAGGGTGACCAAAAAATC
GSTT1	FW: TTCCTTACTGGTCCTCACATCTC RW: TCACCGGATCATGGCCAGCA
GSTM1	FW: GAACTCCCTGAAAAGCTAAAGC RW: GTTGGGCTCAAATATACGGTGG
Alu	FW: GCCCTCTGCTAACAAGTCCTAC RW: GCCCTAAAAAGAAAATCGCCAAT

D-Loop

The mtDNA d-loop region was amplified by PCR using forward 5′-AAGACTCGGCAGCATCTCCACACCATTAGCACCCAAAGCT-3′ and reverse 5′-GCGATCGTCACTGTTCTCCACTGTTAAAAGTGCATACCGCCA-3′ as shown in Table 3. Amplification was performed on VeritiDx, a thermal cycler from Thermo Fisher Scientific using. Primers were used at a final concentration of 200 nM and dNTPs at 0.2 mM; MgCl2 was used at a final concentration of 1.5 mM. An enzyme blend of Platinum taq (Invitrogen, Cat. No. 10966034) and Hotstar taq (Qiagen, Cat. No. 1007837) was used. The reaction mixture was heated to 95 °C for 15 min, followed by 35 cycles each consisting of 1 min denaturation at 95 °C, 1 min annealing at 60 °C, 90 sec of extension at 72 °C and a final 10 min extension at 72 °C as shown in Table

4. The amplified product was purified and sequenced by next generation sequencing of targeted amplicon synthesis chemistry (Illumina MiSeq). Bioinformatic analysis was performed using Illumina's cloud computing platform, Basespace.

Table 4: Thermal profile for D-Loop.

Stage	Steps	Temperature (°C)	Time	No. of Cycles
Stage 1	Initial denaturation	95°C	15 minutes	Hold
Stage 2	Denaturation	95°C	1 minutes	35 cycles
	Annealing	60°C	1 minutes	
	Extension	72°C	90 seconds	
Stage 3	Final Extension	72°C	10 minutes	Hold
Stage 4	Hold	4°C	infinite	Hold

CO1

The mtDNA CO1 region was amplified by PCR using forward 5′-AAGACTCGGCAGCATCTCCATCAACAAATCATAAAGATATTGG-3′ and reverse 5′-GCGATCGTCACTGTTCTCCAGGGTGACCAAAAAATC-3′ as shown in Table 3. Amplification was performed on VeritiDx, a thermal cycler from Thermo Fisher Scientific. Primers were used at a final concentration of 200 nM and dNTPs

Table 5. Thermal profile for CO1.

Stage	Steps	Temperature (°C)	Time	No. of Cycles
Stage 1	Initial denaturation	95°C	15 minutes	Hold
Stage 2	Denaturation	95°C	30 seconds	35 cycles
	Annealing	52°C	30 seconds	
	Extension	72°C	90 seconds	
Stage 3	Final Extension	72°C	10 minutes	Hold
Stage 4	Hold	4°C	infinite	Hold

at 0.2 mM; MgCl2 was used at a final concentration of 2.5 mM. An enzyme blend of Platinum taq (Invitrogen, Cat. No. 10966034) and Hotstar taq (Qiagen, Cat. No. 1007837) was used. The reaction mixture was heated to 95 °C for 15 min, followed by 35 cycles each consisting of 30 sec denaturation at 95 °C, 30 sec annealing at 52 °C, 90 sec of extension at 72 °C and a final 10 min extension at 72 °C as shown in Table 5. The amplified product was purified and sequenced by next generation sequencing of targeted amplicon synthesis chemistry (Illumina MiSeq). Bioinformatic analysis was performed using Illumina's cloud computing platform, Basespace.

Glutathione S-transferase

GSTT1 was amplified by PCR using primers as forward 5′-TTCCTTACTGGTCCTCACATCTC-3′ and reverse 5′-TCACCGGATCATGGCC AGCA -3′; GSTM1 primers as forward 5′-GAACTCCCTGAAAAGCTAAAGC-3′ and reverse 5′-GTTGGGCTCAAATATACGGTGG-3′; and a positive control Alu as forward 5′-GCCCTCTGCTAACAAGTCCTAC-3′ and reverse 5′-GCCCTAAAAAGA AAATCGCCAAT-3′ as shown in Table 3. The PCR conditions for the GSTM, GSTT, and the control gene (*Alu*) were standardized and analysis for these genes was performed in a multiplex PCR according to previously described methods (Arand *et al.,* 1996). The PCR was performed using the QuantiTect 2x mastermix (Qiagen Cat. No. 204543), 200 nM each primer, and approximately 50 ng of DNA extracted from the blood samples provided. The reaction mixture was heated to 95 °C for 15 min, followed by 35 cycles each consisting of 20 sec denaturation at 95 °C, 25 sec annealing at 63 °C, 35 sec of extension at 72 °C and a final 10 min extension at 72 °C as shown in Table 6. The total reaction volume was 20µl; 2 µl of the amplified product was analyzed on a 2% agarose gel which resulted in a 215 bp fragment for GSTM1, 480 bp fragment for GSTT1 and 350 bp fragment for albumin gene as an internal control (Egan *et al*., 2004). The absence of the specific GSTM1 and/or GSTT1 fragments specify the parallel null genotype, whereas the presence of the albumin

Table 6: Thermal profile for multiplex PCR.

Stage	Steps	Temperature (°C)	Time	No. of Cycles
Stage 1	Initial denaturation	95°C	15 minutes	Hold
Stage 2	Denaturation	95°C	20 seconds	35 cycles
	Annealing	63°C	25seconds	
	Extension	72°C	35 seconds	
Stage 3	Final Extension	72°C	10 minutes	Hold
Stage 4	Hold	4°C	infinite	Hold

gene fragment confirms that the accepted null genotype was not due to PCR failure. A No Template Control (NTC), Positive Control (PC) that was positive for all three genes, and FlashGel marker (Lonza) was run with each batch of samples.

4.5. DNA Analysis

D-Loop and CO1

Sequencing of D-loop and CO1 was obtained using NGS (Illumina MiSeq). Genomic visualization was done using Homo sapiens (Human) GRCh37 hg19 (Feb 2009) as reference on GenomeBrowse 3.0.0 and variant analysis was done using VarSeq 1.5.0 (Golden Helix 1.5.0). Variant pathogenicity was done using HmtVar. HmtVar is a manually curated database offering variability and pathogenicity information about mtDNA variants. Data were gathered from HmtVar's twin database HmtDB, and further integrated with pathogenicity predictions as well as additional information from several online resources focused on mtDNA, such as MITOMAP, 1000 Genomes Project, MutPred, SNPs and GO and many others. Sequences were also evaluated against the Mitomaster (https://www.mitomap.org/) for variants and haplogroup. Mitomaster uses HaploGrep2 with Phylotree 17 for haplogroup determination NC_012920.1.

Phylogenetic analysis was performed using the sequences of d-loop and CO1 of all the samples, an outgroup sequence (accession no NC_012920.1) was selected

and retrieved from NCBI. All the sequences were aligned using clustalW implemented in Mega 11 (Tamura *et al.*, 2021) and phylogenetic tree was build using maximum likelihood (ML) and neighbourhood joining (NJ), default parameter was used. Phylip file was generated using ALTER Alignment Transformation Environment (Glez-Peña *et al*., 2010) for RaxML analysis. Maximum likelihood of the sequences was then generated using the program raxmlGUI 2.0 (Edler *et al*., 2021), bootstrap value was set at 500, substitution matrix and substitution rates was set at GTR and GAMMA.

Glutathione S-transferase

The presence of GSTT1 and GSTM1 fragments confirmed that the sample was of normal genotype. Whereas the absence of the GSTT1 and GSTM1 fragments confirmed that the sample was of null genotype. GSTT1 and GSTM1 were combined as having no deletions, wild type GSTT1 with null GSTM1, null GSTT1 with wild type GSTM1 or having homozygous deletions.

4.6. Statistical Methods

Risk factors were compared between cases and controls by constructing frequency distribution using IBM SPSS Statistics 20.0. To test the significance of differences between two groups, Pearson's chi-square test was adopted. Variables were further categorized into 8 groups - demography, dietary habits, behavioral habits, environmental exposure, medical history, menstrual history, reproductive history and family history. Variables with a *P* value less than 0.05 were observed as relevant to show significant differences between case and control groups.

To examine the causal effects of different risk factors, univariate and multivariate logistic regression were used. Odd ratio (ORs) with 95% confidence intervals (CIs) of case and control groups were calculated from these regression estimates to interpret the severity of each factor. In addition to this, survival analysis of the cases was also conducted using Cox-regression model. Subjects with missing values in any of the

variables in a regression model were excluded from the analysis. Variables in which when both case and control have the same value were also excluded (Yu *et al*., 2012).

Statistical test was also performed on a few factors based on menopausal status. A woman was considered postmenopausal: if she had undergone bilateral oophorectomy; if she affirmed that her menstruations had ceased for at least 6 consecutive months before diagnosis of breast cancer or prior to interview for control; if the above information were unavailable or inconclusive, 55 years and above was considered as postmenopausal for the study. To facilitate comparison, variables were categorized in two, never and ever or low and high. To examine the causal effects of the selected factor, univariate and multivariate logistic regression were used. BMI (kg/m^2) were dichotomized based on the median values (22.49) of controls.

Association of risk factors with the selected genes was also performed. To calculate the population characteristics and test differences between molecular study group chi-square was used where appropriate. Variables were also categorized into two where possible. To examine the causal effects of different factors, univariate and multivariate logistic regression were used. Odd ratio (ORs) with 95% confidence intervals (CIs) of case and control groups were calculated from these regression estimates to interpret the severity of each factor. Normal or wild type GSTT1 and GSTM1 genotypes and having no deletion in GSTT1 and GSTM1 were used as referent. The adjusted ORs were evaluated by including the potential confounders such as age, residence, educational qualification and parity.

5. Summary of Results

- From 2003 – 2014 there was approximately 60 cases recorded in a year. According to Mizoram State Cancer Hospital, the oldest recorded case was in 1998 and no trace of earlier records. The number of sporadic cases to familial cases was 309 to 54 which was roughly 85:15, the ratio found worldwide. Within a span of 17 years, out of 758 registered cases 9 were male which is higher than the world data of 1 in 100.
- Incidence is more closely associated to age than to any other risk factor, it increases rapidly during the reproductive years and then more slowly after 50 years of age in this study. The highest incidence was observed in ages between 40 and 49 years.
- The number of Hmar and Lusei tribes are the highest in both the case and control groups, from both the paternal and maternal side. Even though there are no records for population based on tribes, it seems that these two major tribes are having the highest population among the Mizos.
- Strong association with risk was observed for smoking, frequent use of fermented pork fat, having hypertension, late age at menopause and at first childbirth.
- Among the premenopausal groups, heavy physical activity was observed to increase risk. A person can be physically active and yet spend a substantial amount of time being sedentary. From this study it is clear that a typical Mizo woman does not have much leisure time. They may spend a substantial amount of time being sedentary but comparatively they are physically active. If further study could be conducted for association of physical activity in detail, the results might be beneficial. Among the postmenopausal groups, association was observed with being overweight or obese, smoking and frequent use of sodium bicarbonate.
- A follow-up after 5 years showed that survival was influenced by three important factors - they had better survival if they do not smoke, had a lifestyle with moderate physical activity and shorter lifetime cumulative number of menstrual cycles.

- Further studies are required to identify all possible etiologic agents which was not covered in this study. It would be beneficial if further research could be conducted on some of the risk factors identified in this study.
- Affected individuals may associate their malignancy with foods perceived to be poor in nutritional value and over report them relative to unaffected controls. If a prospective study could also be conducted especially on dietary habits, stronger conclusions could have been made on some of the unique regional habits.
- Association of Glutathione s-transferase (M1 & T1) with risk factors does not have any relevant significance in this study group.
- Significant mtDNA mutations in this study either had predisposition to breast cancer or probably damaging or deleterious variant as established from other published data. Interaction of these mutations with behavioural habits reveal significant risk with use of betel quid, tobacco, tuibur, smoking, reduced sleep duration and sleep efficiency. Interaction of miscarriages and older age of mother with these mutations also had significant association.
- The established factors that are modifiable observed in this study are late age at first child birth and being obese or overweight after menopause; the non-modifiable are age and late age at menopause; the suspected are smoking cigarette and hypertension; the unclear but suspected found specific to this region is frequent use of soda bicarbonate and fermented pork fat.

Lalhmachhuani

June 27, 2023