Genetic polymorphisms in the folate metabolizing pathway and breast cancer survival.

Nandita Das

University of Louisville
GENETIC POLYMORPHISMS IN THE FOLATE METABOLIZING PATHWAY
AND BREAST CANCER SURVIVAL

By

Nandita Das
M.A., Tezpur Central University, 2001
M.A., University of Louisville, 2006

A Dissertation
Submitted to the Faculty of the
School of Public Health and Information Sciences of the University of Louisville
In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

Department of Epidemiology and Population Health
University of Louisville
Louisville, Kentucky

May 2012
Copyright 2012 by Nandita Das

All rights reserved
GENETIC POLYMORPHISMS IN THE FOLATE METABOLIZING PATHWAY 
AND BREAST CANCER SURVIVAL

By

Nandita Das
M.A., Tezpur Central University, 2001
M.A., University of Louisville, 2006

A Dissertation Approved

April 26, 2012

By the following Dissertation Committee:

Richard Baumgartner, Dissertation Director

__________________________
Richard A. Kerber

__________________________
Kathy B. Baumgartner

__________________________
Shesh N. Rai

__________________________
Cornelia Ulrich
DEDICATION

This dissertation is dedicated to my parents
Late Lakshmi Prasad Lahkar and Mrs. Alaka Lahkar
and my beloved husband
Mr. Ratan Kumar Das
who has always been my source of inspiration.
ACKNOWLEDGEMENTS

First of all I would like to express my sincere gratitude to my mentor Dr. Richard N. Baumgartner, who has been an endless source of knowledge and guidance. He has given me the opportunity to be a graduate assistant for the department of Epidemiology and be a part of the Healthy Eating Activity and Life (HEAL) study which he along with other experts in the field initiated approximately fifteen years ago. He has not only guided me throughout this project, but also coached me every step of the way during my entire doctoral degree program. To me Dr. Rick Baumgartner is a true epitome of a “Guru”, a figure often associated with Brahma in Indian culture. The skills that I acquired from him during the course of the past few years have not only made me a better Epidemiologist but also shaped my personality in a way that will leave lasting impressions on my life.

If Dr. Rick Baumgartner reminds me of Lord Brahma, Dr. Kathy B. Baumgartner makes me think of Saraswati, the Hindu goddess of knowledge and wisdom. She along with Dr. Rick Baumgartner recruited me as a Susan G. Komen Trainee. Being a part of her Long Term Quality of Life (LTQOL) study, which is a fifteen years follow-up study of the New Mexico Women’s Health Study (NMWHS), she helped me acquire first-hand experience on various aspects of Epidemiological research. Throughout the program, she identified my strengths and coached me to improve upon them and pointed at my flaws and
suggested ways to rectify them. Being a woman, someday I aspire to walk on her footsteps and be a well-rounded individual and be a superb mentor like her.

I sincerely appreciate Dr. Richard A. Kerber for teaching me the ABCs of Genetic Epidemiology. He instilled in me a sense of confidence in exploring the highly complex and ever expanding world of Genetic Epidemiology. Without his guidance I would not have been able to initiate this project. I would also like to express my thankfulness to Dr. Cornelia Ulrich, who is a world renowned expert in the field of Nutritional Epidemiology and spent most of her career studying the folate pathway. Her expertise was extremely helpful for this project. I would like to extend my gratitude to Dr. Shesh N. Rai for helping me with the statistical aspects of this project.

I would also like to thank the entire panel of experts in the HEAL team for allowing me to use the data. A very special thanks to the project officer, Dr. Rachel Ballard-Barbash, and the co-PIs of the HEAL study, Dr. Anne McTiernan, and Dr. Leslie Bernstein for their constant support throughout this project. I would like to thank Dr. Marian L. Neuhouser for her valuable insight on the dietary intakes data. Two other valuable members, Anita Ambs, and Todd Gibson were always helpful to me throughout the project. I really appreciate them both.

I would like to express my gratitude to all the participants of the HEAL study who graciously contributed their time, shared their personal and private experiences during the interviews at various time points, and contributed their
biological specimens for the cause of breast cancer science. Special thanks to Susan G. Komen Foundation for providing the funding for this project.

I would also like to extend my thanks to my two fellow Susan G. Komen scholars and LTQOL study team mates, Dr. Avonne C. Connor and Stephanie Denkhoff, for helping me with the lab work. Special thanks to Paula K. Bossmeyer, Jason Banta, Tammi Thomas and the entire School of Public Health for helping me throughout the PhD program and allowing me to be a part of the community.

I also take this opportunity to thank my two brothers who always stood by my side in every step of my life. After my father passed away almost two decades ago, I realized that he left behind for me two fathers to follow my dreams and learn to be a better human being in the process. I am sure my father will continue to shower his blessings on me and the rest of the family. Words fall short while trying to express my gratitude to my mother who has always been a source of inspiration for me. She may live across the globe but always a phone call away from me when I need her. Last but not least I would like to express my thanks to my beloved husband, Ratan, for always lifting my spirits with his sense of humor. He brings out the best in me by keeping me focused and on schedule. I feel blessed to be his wife.
ABSTRACT

Genetic polymorphisms in the folate metabolizing pathway and breast cancer survival

Nandita Das

April 26, 2012

Genes in folate metabolizing pathway are responsible for modulating key enzymes in folate metabolisms necessary for DNA synthesis and repair and influence DNA methylation. Irregularity in DNA methylation may modify tumor phenotype response to chemotherapy and abnormal DNA synthesis, and repair may lead to carcinogenesis. This dissertation project tested the hypothesis that the key genes in the folate metabolizing pathway, namely Methylenetetrahydrofolate reductase (MTHFR), betaine-homocysteine-methyltransferase (BHMT), 5-methyltetrahydrofolate-homocysteine (MTR), 5-methyltetrahydrofolate-homocysteine methyltransferase reductase (MTRR), and serine hydroxymethyltransferase (SHMT1) and folate receptor 1 (FOLR1) influences all-cause and breast cancer-specific mortality and the association between the genes in the folate metabolizing pathway and breast cancer-specific survival is modified by chemotherapy and estrogen receptor (ER) status. Gene-gene interactions among the genes in the folate metabolizing pathway were also tested in this study. Data for this study were obtained from the New Mexico site of the Health, Eating, Activity, and Lifestyle (HEAL) Study for 446 incident
primary breast cancer cases (stage 0 - IIIA). Over 11.46 years of follow-up, there were 67 deaths; 22 due to breast cancer. In the Cox proportional hazards analysis, women who carried at least one variant allele for MTRR G66A polymorphism had statistically significant inverse associations with all-cause (Hazard Ratio [HR]: 0.51; 95% confidence interval [CI]: 0.3-0.8; p-value <0.01) and breast cancer-specific mortality (HR: 0.40; 95% CI: 0.2-0.9; p-value <0.01). There were no statistically significant interactions with chemotherapy or ER status. Synergistic multiplicative interaction was observed between BHMT A742G and FOLR1 G606A polymorphisms in predicting breast cancer-specific mortality (p for interaction: 0.03). In this sample, there was antagonistic multiplicative interaction between MTRR G66A and MTHFR C677T polymorphisms in predicting breast cancer-specific mortality (p for interaction: 0.08). These results should be further explored in larger cohorts and in clinical populations.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGEMENTS</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
</tbody>
</table>

INTRODUCTION..................................................................... 1

  Study Rationale....................................................... 1
  Objectives....................................................................... 3

REVIEW OF THE LITERATURE.................................................. 6

  Background..................................................................... 6
  Primary Independent Variables.................................. 6
  Covariates of Interest............................................... 21

METHODS AND METHODOLOGY.................................................... 26

  Introduction............................................................. 26
  Study Population...................................................... 26
  Hypothesis................................................................. 30
  Statistical Plan......................................................... 32
  Operational Hypothesis............................................... 38
  Power Calculations.................................................... 40
  Genotyping..................................................................... 48

PRELIMINARY STUDY: MTHFR POLYMORPHISMS.......................... 51
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Genes in the one-carbon metabolizing pathway along with their respective physical location and nucleotide change</td>
<td>9</td>
</tr>
<tr>
<td>2. Values of f and g for dominant and recessive models</td>
<td>43</td>
</tr>
<tr>
<td>3. Power (two and one-tailed) to test association between genotype and breast cancer mortality for the main effects of the dominant models using all three HEAL sites</td>
<td>44</td>
</tr>
<tr>
<td>4. Power (two and one-tailed) to test association between genotype and breast cancer mortality for the dominant models stratified by use of chemotherapy using all three HEAL sites</td>
<td>45</td>
</tr>
<tr>
<td>5. Power (two and one-tailed) to test association between genotype and breast cancer mortality for the main effects of the dominant models using samples from New Mexico site</td>
<td>46</td>
</tr>
<tr>
<td>6. Power (two and one-tailed) to test association between genotype and breast cancer mortality for the dominant models stratified by use of chemotherapy using samples from New Mexico site</td>
<td>47</td>
</tr>
<tr>
<td>7. Calculations for Hardy Weinberg Equilibrium for the SNPs investigated in this study</td>
<td>50</td>
</tr>
<tr>
<td>8. Demographic and medical characteristics of breast cancer survivors in the Health, Eating, Activity, and Lifestyle study</td>
<td>54</td>
</tr>
</tbody>
</table>
9. Crude Hazard Ratios for any cause mortality and breast cancer-specific mortality for study variables ........................................ 55

10. Prognosis: stratified by chemotherapy and genotype showing breast cancer mortality ........................................................................................................ 56

11. Interaction of genotype with chemotherapy on breast cancer-specific mortality .................................................................................................................. 57

12. Demographic and medical characteristics of breast cancer survivors for the New Mexico site in the Health, Eating, Activity, and Lifestyle Study ................................................................................................................. 68

13. Genotypes of breast cancer survivors for the New Mexico site in the Health, Eating, Activity, and Lifestyle Study ................................................................................................................. 69

14. Age-adjusted hazard ratios and 95% confidence intervals for the associations of polymorphisms in the folate metabolizing pathway and all-cause mortality and breast cancer-specific mortality using data from the New Mexico site of the Health, Eating, Activity, and Lifestyle study ................................................................................................................. 72

15. Age-adjusted breast cancer-specific mortality stratified by genotypes in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study ................................................................................................................. 76

16. Age-adjusted breast cancer-specific mortality stratified by chemotherapy and genotype in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study ................................................................................................................. 82

17. Age-adjusted breast cancer-specific mortality stratified by ER
status and genotype in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study ..................................................... 89

18. Amounts of ingredients required to make reaction mixtures ........ 123
<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Simplified version of the folate metabolizing pathway</td>
<td>8</td>
</tr>
<tr>
<td>2. Diagrammatic representation of chromosome 1 showing <em>MTHFR</em></td>
<td>11</td>
</tr>
<tr>
<td>3. Diagrammatic representation of chromosome 1 showing <em>MTR</em></td>
<td>13</td>
</tr>
<tr>
<td>4. Diagrammatic representation of chromosome 5 showing <em>MTRR</em></td>
<td>14</td>
</tr>
<tr>
<td>5. Diagrammatic representation of chromosome 5 showing <em>BHMT</em></td>
<td>15</td>
</tr>
<tr>
<td>6. Diagrammatic representation of chromosome 17 showing <em>SHMT1</em></td>
<td>16</td>
</tr>
<tr>
<td>7. Diagrammatic representation of chromosome 11 showing <em>FOLR1</em></td>
<td>17</td>
</tr>
<tr>
<td>8. Cohort definition for the New Mexico site of the Health, Eating, Activity, and Lifestyle (HEAL) Study and the number of events during 11.46 years of follow-up</td>
<td>29</td>
</tr>
<tr>
<td>9. Health, Eating, Activity, and Lifestyle (HEAL) cohort definition and number of events during 11.46 years of follow-up stratified by chemotherapy</td>
<td>53</td>
</tr>
<tr>
<td>10. Kaplan-Meier curve for ten years breast cancer-specific mortality by chemotherapy for carriers of CC genotype (C677T polymorphism) showing number of months from diagnosis to breast cancer death using data from Health, Eating, Activity, and Lifestyle study</td>
<td>60</td>
</tr>
</tbody>
</table>
11. Kaplan-Meier curve for ten years breast cancer-specific mortality by chemotherapy for carriers of CT/TT genotype (C677T polymorphism) showing number of months from diagnosis to breast cancer death using data from Health, Eating, Activity, and Life-style study...........60

12. Kaplan-Meier curve for ten years breast cancer-specific mortality for carriers of A1298C polymorphism showing number of months from diagnosis to breast cancer death using data from Health, Eating, Activity, and Life-style study........................................61

13. Patterns of recombination showing change in C677T and A1298C polymorphisms.................................................................65

14. Kaplan-Meier survival curves among carriers of the BHMT A742G, AA genotype stratified by FOLR1 G606A polymorphism, HEAL study, New Mexico...............................................................................................78

15. Kaplan-Meier survival curves among carriers of the BHMT A742G, AG/GG genotype stratified by FOLR1 G606A polymorphism, HEAL study, New Mexico...............................................................................................78

16. Kaplan-Meier survival curves among carriers of the MTRR G66A, GG genotype stratified by MTHFR C677T polymorphism, HEAL study, New Mexico...........................................................80

17. Kaplan-Meier survival curves among carriers of the MTRR G66A, GA/AA genotype stratified by MTHFR C677T polymorphism, HEAL study, New Mexico...........................................................80
18. Kaplan-Meier survival curves among carriers of the *BHMT* A742G, AA genotype stratified by chemotherapy, HEAL study, New Mexico............................................................. 85

19. Kaplan-Meier survival curves among carriers of the *BHMT* A742G, AG/GG genotype stratified by chemotherapy, HEAL study, New Mexico............................................................. 85

20. Kaplan-Meier survival curves among carriers of the *SHMT* G1420A, GG genotype stratified by chemotherapy, HEAL study, New Mexico............................................................. 87

21. Kaplan-Meier survival curves among carriers of the *SHMT* G1420A, GA/AA genotype stratified by chemotherapy, HEAL study, New Mexico............................................................. 87

22. Kaplan-Meier survival curves among carriers of the *MTRR* G66A, GG genotype stratified by ER-status, HEAL study, New Mexico............................................................. 92

23. Kaplan-Meier survival curves among carriers of the *MTRR* G66A, GA/AA genotype stratified by ER-status, HEAL study, New Mexico............................................................. 92

24. Kaplan-Meier survival curves among carriers of the *MTR* A2756G, AA genotype stratified by ER-status, HEAL study, New Mexico............................................................. 94

25. Kaplan-Meier survival curves among carriers of the *MTR* A2756G, AG/GG genotype stratified by ER-status, HEAL study, New Mexico............................................................. 94
26. Animated representation of the 5' nuclease TaqMan assay

process..................................................................................118
CHAPTER I
INTRODUCTION

STUDY RATIONALE

According to the American Cancer Society’s 2012 report\(^1\), breast cancer is the second leading cause of cancer death among women in the US, with over forty thousand expected deaths in 2012 alone\(^1\). Studies show that folate deficiency may be associated with breast cancer risk and may also influence breast cancer-specific survival\(^2\textsuperscript{4}\). Folate plays an important role in DNA methylation\(^2\). It helps in the synthesis of purines, thymidylate, glycine, and methionine thus serving an important role in nucleotide synthesis\(^2\). Its deficiency could, therefore, lead to abnormal DNA methylation. Abnormal DNA synthesis, and repair may lead to carcinogenesis and irregularity in DNA methylation may modify response to chemotherapy\(^5\textsuperscript{6}\). Genes in folate- or one-carbon metabolizing pathway are responsible for modulating key enzymes in folate metabolisms necessary for DNA synthesis and repair and influence DNA methylation\(^2\textsuperscript{7}\textsuperscript{8}\). Methylenetetrahydrofolate reductase (\textit{MTHFR}), betaine-homocysteine-methyltransferase (\textit{BHMT}), 5-methyltetrahydrofolate-homocysteine (\textit{MTR}), 5-methyltetrahydrofolate-homocysteine methyltransferase reductase (\textit{MTRR}), and serine hydroxymethyltransferase (\textit{SHMT1}) are some of the key genes in the folate metabolizing pathway responsible for the enzymes in folate metabolism. \textit{MTHFR}, \textit{MTR}, \textit{MTRR}, \textit{BHMT} and \textit{SHMT1} act as catalysts in
the conversion of 5, 10-methylTHF to 5-methyleneTHF, a donor for the remethylation of homocysteine to methionine, the precursor for S-adenosylmethionine (SAM)\(^9\). Folate receptor 1 (FOLR1), which lies outside the one-carbon metabolizing pathway, modulates cellular folate availability, which may indirectly influence intracellular metabolism and subsequently DNA methylation, and synthesis. Also taken together these genes regulate folate availability and metabolism. Polymorphic variations in these genes, in the presence of folate deficiency, could lead to abnormal DNA methylation, and synthesis\(^10\).

Several studies have investigated the association of \textit{MTHFR} polymorphisms, dietary folate, and breast cancer risk\(^{11-17}\). Three studies have so far investigated the association between \textit{MTHFR} polymorphisms and breast cancer survival\(^{12,18,19}\) of which only one study, the Long Island Breast Cancer Study\(^{19}\) has investigated the associations of the other genes in the folate metabolizing pathway with breast cancer survival. The Long Island Breast Cancer Study\(^{19}\) with an average follow-up of 5.6 years and a sample size of 1,508 women diagnosed with both \textit{in situ} and invasive breast tumors, reported an association between the \textit{MTHFR} (C677T) polymorphism and breast cancer mortality. In their sample, estrogen receptor (ER) status significantly modified the effect of the association between C677T breast cancer-specific survival. The Shanghai Breast Cancer Study\(^{20}\) did not detect any significant association between \textit{MTHFR} polymorphisms and five year survival. C677T genotype non-significantly increased breast cancer-specific mortality among women diagnosed
with late stage (stage III-IV) tumor. Martin et al\textsuperscript{18} observed similar inverse association between C677T and breast cancer-specific mortality. They also investigated the effects of ER status and race on the association between \textit{MTHFR} polymorphisms and breast cancer-specific mortality. In their sample, women with ER negative tumors had worse survival for both C677T and A1298C and race (Caucasian vs. African American) modified the effect of \textit{MTHFR} polymorphisms on breast cancer-specific mortality.

**Study Objectives**

The overall objective of this dissertation is to investigate the association between genetic variants in the folate metabolizing pathway and the genes responsible for maintaining blood folate level and breast cancer-specific survival and to examine the effects of chemotherapy use, ER status, and folate intake on these associations. Additionally, gene-gene interaction was evaluated.

The specific aims of this dissertation include:

1. To investigate the associations between genetic variants in the folate metabolizing pathway (\textit{BHMT}, \textit{SHMT1}, \textit{MTR} and \textit{MTRR}) and the gene responsible for maintaining blood folate level (\textit{FOLR1}) in relation to breast cancer-specific survival, adjusted for age at diagnosis, ethnicity, stage, chemotherapy, ER-status, folate and B-complex vitamin, alcohol and energy intake.

Genes in folate metabolizing pathway are responsible for modulating key enzymes in folate metabolisms necessary for DNA synthesis and repair. Several studies have investigated the association of \textit{MTHFR} polymorphisms and breast
cancer risk particularly with respect to the modifying effect of \textit{MTHFR} polymorphisms on the association between folate intake and breast cancer risk. Associations between genes in the folate metabolizing pathway and risk of colorectal cancer and adenomas have also been investigated. Studies have also investigated the association between \textit{MTHFR} polymorphisms and breast cancer survival but only one cohort study, the Long Island Breast Cancer Study, has so far investigated the association between other genes in the folate metabolizing pathway and breast cancer mortality.

2. To investigate if breast cancer survival associated with each of these genes differs when other genes in the same pathway are present (gene-gene interaction).

Since the enzymes in a pathway may interact with each other, it is important to investigate the interactions among genes in the folate metabolizing pathway on breast cancer-specific survival. Studies have not yet investigated the effect of gene-gene interactions on folate metabolizing pathway in predicting breast cancer survival. Hence this dissertation project, sought to test if the presence of variants of one gene influences the associations of variants in other genes in the folate metabolizing pathway with breast cancer survival.

3. To investigate if breast cancer survival associated with each of these genes differs with use of (a) chemotherapy, and (b) ER status.

In breast cancer cell lines, mice models and breast cancer patients, \textit{MTHFR} polymorphisms are reported to alter the chemo-sensitivity of commonly used chemotherapeutic agents in breast cancer treatment but few studies have
investigated the effects of chemotherapy on the associations of genes in the folate metabolizing pathway with breast cancer-specific survival. ER status is a widely used prognostic marker and indicator of breast cancer treatment\textsuperscript{28}. Since the Martin et al\textsuperscript{18} and Xu et al\textsuperscript{19} have investigated the effects of ER status on the association of genes in the folate metabolizing and breast cancer survival, the current study is an attempt to replicate and extend their findings.

In short, this dissertation will help to enrich the current literature on the folate metabolizing pathway and breast cancer survival. Since ER status is a widely accepted prognostic marker in suggesting breast cancer treatment, its effects on the genes in the pathway in predicting breast cancer survival rates may help clinicians in future to decide the types of treatment suggested to women with a certain genetic make-up. The findings of current study may also help to clarify research to chemotherapy. Future studies should include cohorts and randomized clinical trials to further explore the findings of the current study.
CHAPTER II

LITERATURE REVIEW

INTRODUCTION

Among the U.S. women breast cancer is the second leading cause of cancer death, following lung cancer\(^1\). Not including skin cancer, it is also the most commonly diagnosed form of cancer among women. According to the estimates of the American Cancer Society, the lifetime risk for being diagnosed with breast cancer in the United States is approximately 1 in 8 and the probability of dying from it is about 1 in 35\(^{29}\). During 2012 alone, an estimated 226,870 new cases of invasive and 63,300 new cases of \textit{in situ} breast cancer will be diagnosed in the US\(^1\). In the past few decades the rates of survival following breast cancer diagnosis has significantly increased mostly due to improvement in treatment and early detection\(^{29}\).

PRIMARY INDEPENDENT VARIABLES

1. **Folate metabolizing Pathway**

   Folate deficiency may be associated with breast cancer risk and breast cancer-specific survival\(^{2-4}\). Folate helps in the synthesis of purines, thymidylate, glycine, and methionine thus serving an important role in nucleotide synthesis and intracellular methylation\(^4\). Its deficiency could, therefore, lead to abnormal DNA methylation. Genes in the folate metabolizing pathway like \textit{MTHFR} (C677T;
rs1801133 and A1298C; rs1801131), \textit{BHMT} (A742G; rs3733890), \textit{MTR} (A2756; rs1805087), \textit{MTRR} (G66A; rs1801394), \textit{SHMT1} (G1420A; rs1979277) and \textit{FOLR1} (G606A; rs3796517) encode for key enzymes in folate metabolism, which is essential in the process of DNA methylation, synthesis and repair \cite{27,8}. DNA methylation involves addition of the methyl group to the 5th position of the cytosine pyrimidine ring which helps regulates cell development and differentiation. Methylation occurs in CpG (cytosine-phosphate-guanine) islands rich in C and G, upstream of the promoter region and is facilitated by a group of enzymes called the DNA methyltransferases. Figure 1 provides a diagrammatic representation of the pathway. To investigate the influence of these genes, one of their single nucleotide polymorphisms (SNP) which has been shown in the literature to be associated with other cancers in humans was selected. Table 1 lists the genes in the folate metabolizing pathway along which are of interest in this dissertation project, their location, function, sequence showing nucleotide change, allele frequency estimates and the populations from which the frequency estimates were estimated.
Figure 1: Simplified version of the folate metabolizing pathway; figure modified from ref\textsuperscript{11,12}
Table 1 Genes in the one-carbon metabolizing pathway along with their respective physical location and nucleotide change.  

<table>
<thead>
<tr>
<th>Gene/Polymorphs</th>
<th>Function</th>
<th>Gene Location</th>
<th>Sequence</th>
<th>Allele Freq (%) Population*</th>
</tr>
</thead>
</table>
| **MTHFR**  
(Methylene tetrahydrofolate reductase)  
rs1801133 (C677T) | Catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which acts as a donor for the remethylation of homocysteine to methionine.  
1p36.22  
(11,785,525-11,788,057 bp) | AGGTGTCTGCGGGAG[CT]  
CGATTTGATCATCAC | EU  
23-31 |  
AS  
25-51 |  
AA  
8-13 |  
SSA  
10-36 |
| **MTHFR**  
(Methylene tetrahydrofolate reductase)  
rs1801131 (A1298C) | Catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which acts as a donor for the remethylation of homocysteine to methionine.  
1p36.22  
(11,785,525-11,788,057 bp) | AGCTGACCAGTGAAG[A/C]  
AAATGCTTTGAAGT | EU  
34-50 |  
AS  
14-25 |  
AA  
12-13 |  
SSA  
10-12 |
| **MTR**  
(5-methyltetrahydrofolate-homocysteine)  
rs1805087 (A2756G/D9 19G) | Catalyzes the final step in methionine synthesis.  
1q43  
(236958581-237067280 bp) | GGAAGAATATGAAGATA  
TTAGACAGG[A/G]CCATT  
ATGAGTCTCTCAAGGTAAGT | EU  
16-50 |  
AS  
7-23 |  
AA  
23-42 |  
SSA  
30-32 |

*EU (European); AS (Asian); SSA (sub Saharan African); AA (African American); ** Location type: Missense
Table 1 Continued: Genes in the one-carbon metabolizing pathway along with their respective physical location and nucleotide change

<table>
<thead>
<tr>
<th>Gene/Polymerphs</th>
<th>Function</th>
<th>Gene Location</th>
<th>Sequence</th>
<th>Allele Freq (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MTRR</strong> (5-methyltetrahydrofolate homocysteine methyltransferase reductase)</td>
<td>Generate functional methionine synthase which serves as catalyst during methionine synthesis, essential for protein synthesis and OCM (^{24}).</td>
<td>5p15.31 (7869217-7901233 bp)</td>
<td>CAGGCAAGAGGCGATCG CAGAAAGAAT(A</td>
<td>G</td>
</tr>
<tr>
<td>rs1801394 (A66G/I22M)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BHMT</strong> (betaine-homocysteine methyltransferase)</td>
<td>Catalyzes the conversion of betaine and homocysteine to dimethylglycine and methionine, respectively (^{25}).</td>
<td>5q13.1-q15 (78,407,604-78,428,112 bp)</td>
<td>CATGAAGGAGGGCTTG GAGGCTGCC[A</td>
<td>G</td>
</tr>
<tr>
<td>rs3733890 (G742A)</td>
<td></td>
<td></td>
<td></td>
<td>AS 21-36</td>
</tr>
<tr>
<td><strong>SHMT1</strong> (serine hydroxymethyltransferase 1)</td>
<td>Catalyzes the reversible conversion of serine and tetrahydrofolate to glycine and 5,10-methylene tetrahydrofolate which donates one carbon units for methionine, thymidylate, and purines synthesis (^{26}).</td>
<td>17p11.2 (18243719-18266856 bp)</td>
<td>GTCAAGGAGGCGAGGC ACAAGGAAAGA[A</td>
<td>G</td>
</tr>
<tr>
<td>rs1979277 (C1420T)</td>
<td></td>
<td></td>
<td></td>
<td>AS 2-8</td>
</tr>
<tr>
<td><strong>FOLR1</strong> (folate receptor 1)</td>
<td>Binds with folic acid and its derivatives and transport 5-methylTHF into cells (^{27}).</td>
<td>11q13.3-q14.1 (71903173-71907366 bp)</td>
<td>AAATTAAGTGGATGGAA GAGAAACTA[A</td>
<td>G</td>
</tr>
<tr>
<td>rs3796517 (G606A)</td>
<td></td>
<td></td>
<td></td>
<td>AS 15-22</td>
</tr>
</tbody>
</table>

*EU (European); AS (Asian); SSA (sub Saharan African); AA (African American); ** Location type: Missense
Following is a brief overview of the genes in the folate metabolizing pathway that are investigated in this dissertation:

**MTHFR (C677T and A1298C):**

*MTHFR* is an enzyme that is coded by the gene with the same name and symbol located at chromosome 1, p36.22 between 11,785,525 and 11,788,057 bp. It is a key enzyme in folate metabolism, which is essential in the process of DNA methylation, synthesis and repair. Figure 2 below provides a diagrammatic representation of chromosome 1 showing the location of *MTHFR* gene.

![Diagram of chromosome 1 showing the location of MTHFR gene](image)

**Figure 2: Diagrammatic representation of chromosome 1 showing MTHFR**

Studies on genetic polymorphisms of *MTHFR* have identified two common single nucleotide polymorphisms (SNP): C677T (rs1801133) and A1298C (rs1801131) with population minor allele frequencies of 32% and 23% respectively.

Results from the Long Island Breast Cancer Study suggested significant reduction in breast cancer-specific mortality rates among C677T allele carriers compared to non-carriers (HR= 0.58; 95% CI: 0.38 - 0.89). They also reported that estrogen receptor-progesterone receptor (ER/PR) status modified the effect of C677T polymorphism breast cancer mortality (p = 0.05). No significant association was observed between A1298C genotype and breast cancer
mortality. The Shanghai BC Study\textsuperscript{20} did not detect any significant association between \textit{MTHFR} polymorphisms and five year survival. C677T genotype significantly increased the likelihood (HR= 2.97, 95\% CI: 1.10, 7.98) of long term mortality among patients diagnosed with late stage (stage III-IV) disease. Martin et al\textsuperscript{18} observed similar inverse association between C677T and breast cancer-specific mortality. They also investigated the effects of ER status and race on the association between \textit{MTHFR} polymorphisms and breast cancer-specific mortality. In their sample, women with ER negative tumors had worse survival (C677T: HR: 0.36; 95\% CI: 0.12-1.35; A1298C: HR: 2.70; 95\% CI: 1.17-6.23) and there was significant interaction between race (Caucasian vs African American) and breast cancer mortality (C677T, p for interaction: 0.03; A1298C, p for interaction: 0.09). These studies, however, did not investigate potential effect modification of the association between \textit{MTHFR} polymorphisms and breast cancer mortality by the use of chemotherapy. During the preliminary phases of this study, associations between \textit{MTHFR} polymorphisms and breast cancer mortality, particularly with respect to the modifying effect of chemotherapy on these associations were investigated using data from all three sites of the Health, Eating, Activity and Lifestyle (HEAL) Study (Refer to chapter IV for details related to findings of the preliminary study; unpublished results).

\textit{MTR (A2756G)}:

\textit{MTR} encodes for the enzyme 5-methyltetrahydrofolate-homocysteine methyltransferase, which is a Vitamin B\textsubscript{12} dependent enzyme\textsuperscript{32} (figure 1). It is also known as methionine synthase. It catalyzes the methylation of
homocysteine to methionine and simultaneously converting 5-methylTHF to THF required for nucleotide synthesis\textsuperscript{33}. \textit{MTR} is located in chromosome 1(1q43) (figure 3).

Figure 3: Diagrammatic representation of chromosome 1 showing \textit{MTR}

A2756G (rs1805087) is a common variant of the gene which substitutes the amino acid aspartic acid to glycine at codon 919\textsuperscript{34,35}. Although studies have reported the effects of \textit{MTR} A2756G on plasma homocysteine in humans, they are inconsistent about the levels of homocysteine across genotypes of A2756\textsuperscript{34-36}. Evidence of the association between folate and B complex vitamins and homocysteine levels is limited\textsuperscript{37}. The current literature is also inconsistent about the effects of \textit{MTR} A2756G on breast\textsuperscript{35,38-41} and colon cancer\textsuperscript{22,42} risk, particularly with relation to established risk factors like menopausal status, alcohol intake, and folate intake. There is insufficient research to date on the association of \textit{MTR} and breast cancer mortality.

\textbf{\textit{MTRR (G66A):}}

\textit{MTRR} gene encodes for the enzyme methionine synthase reductase which acts as catalyst in the final step of the conversion of homocysteine to methionine. \textit{MTRR} is located in chromosome 5 (5p15.31) (figure 4).
Figure 4: Diagrammatic representation of chromosome 5 showing \textit{MTRR}

G66A (rs1801394) is a common variant of the gene which substitutes isoleucine with methionine at codon 22\textsuperscript{43} and as the name suggests (reductase), activates methionine synthase, also known as \textit{MTR} by acting as a carbon donor for homocysteine remethylation. Like the \textit{MTR} polymorphism, studies have consistently observed an association between homocysteine levels and \textit{MTRR} G66A but while some studies report elevated levels of homocysteine among carriers of AA genotype\textsuperscript{44,45} others report decrease in homocysteine levels \textsuperscript{46}. Studies also did not find significant associations between folate, B-complex vitamins, and alcohol intake and \textit{MTRR} G66A genotypes. In order to better understand the impact of the variant of \textit{MTRR} G66A in predicting breast cancer survival, research on this polymorphism is due.

\textit{BHMT} (A742G):

\textit{BHMT} encodes for the protein betaine-homocysteine methyltransferase and is located in chromosome 5 (5q13.1-q15) (Figure 5). Unlike \textit{MTR}, \textit{MTRR}, and \textit{MTHFR}, \textit{BHMT} is not involved in the folate-dependent re-methylation process. Instead \textit{BHMT} utilizes betaine, a metabolite of choline, as its carbon donor in the remethylation of homocysteine to methionine (figure 1)\textsuperscript{47}. 
Studies have observed significant correlation between plasma homocysteine and plasma betaine levels and have suggested that plasma betaine is capable of lowering plasma homocysteine levels\(^4^7\). So \textit{BHMT} plays an important role in the re-methylation of homocysteine in the folate metabolizing pathway when the pathway is compromised due to mutations in the \textit{MTRR} or \textit{MTR} polymorphisms or deficiencies in vital nutrients\(^4^7\)\(^4^8\). As for \textit{MTHFR} polymorphisms, studies have investigated the association of \textit{BHMT} A742G with risk for hyperhomocysteinemia\(^4^9\), vascular diseases\(^4^9\)\(^5^0\) and congenital diseases like neural tube defect\(^5^1\), and Downs Syndrome\(^5^2\). Studies have yet to investigate the association between \textit{BHMT} A742G and breast cancer survival.

\textbf{\textit{SHMT1} (C1420T/G1420A)}:

\textit{SHMT1}, sometimes called cytosolic serine hydroxymethyltransferase (\textit{cSHMT1}), encodes for the protein of the same name and is one of the less frequently studied genes in the folate metabolizing pathway. It is located in chromosome 17 (17p11.2) (figure 5).
Figure 6: Diagrammatic representation of chromosome 17 showing SHMT1

It has a functional Leu > Phe polymorphism (C1420T/G1420A; rs1979277). It is vitamin-B_6 dependent enzyme that facilitates SAM synthesis and catalyzes the conversion from 5-methylTHF to 5-10-methyleneTHF. It has been suggested that carriers of the minor alleles of the SHMT1 G1420A polymorphism may decrease production of the 5-10-methylTHF for pyrimidine synthesis. This may reduce levels of DNA methylation through negative feedback which may in turn result in decreased 5-methylTHF production and accumulation of THF, leading to carcinogenesis\(^{53}\). It has been studied in relation to risk for acute lymphocytic leukemia\(^{54}\), malignant lymphoma\(^{55}\), and colorectal cancer\(^{53}\)\(^{56}\). The current research is an attempt to contribute to the literature on the SHMT1 polymorphism and breast cancer mortality.

**FOLR1 (A606G):**

As mentioned already, although outside of the folate metabolizing pathway, the folate receptor α enzyme encoded by the FOLR1 gene is located in chromosome 11 (11q13.3-q14.1) (Figure 6) plays an important role in the uptake of folate in the cell\(^{57}\).
While dietary folate is known to exist in a polyglutamated state, folic acid from fortified foods and supplements is in monoglutamated state and is taken up by the cell differently. Folic acid does not require hydrolysis before absorption.

Once folate is inside the intestinal cells, regardless of poly or mono-glutamated states, various coenzymes convert it to 5-methyl THF prior to release into the bloodstream. Folate receptor α binds with folic acid and its derivatives and helps transport 5-methylTHF for SAM synthesis. Depending on the type of tissue, 5-methyl THF is also taken up by tissues via folate receptor α from the bloodstream\(^\text{57}\). Like SHMT1, FOLR1 G606A (rs3796517) is also one of the less frequently studied polymorphism associated with folate metabolism. Effects of FOLR1 polymorphisms have been investigated along with genes in the folate metabolizing pathway with respect to renal cell\(^\text{58}\) and prostate\(^\text{59}\) cancer risk. But studies have not yet investigated its association with breast cancer survival, so in this dissertation the effects of the variant of the FOLR1 G606A are explored.

2. Chemotherapy:

Biochemical studies have shown that besides being an anti-metabolite of folate, two metabolites of fluoropyrimidine (5-FU) a commonly used chemotherapeutic agent in breast cancer treatment, 5-fluoro-2'-deoxyuridine-5'-triphosphate and 5'-flurouridine-5'-triphosphate, can interfere with the level of
DNA stability and RNA functioning. Further, 5-fluoro-2'-deoxyuridine-5'-monophosphate (5FdUMP), an active metabolite of 5-FU, forms a complex with TS and 5,10-methyleneTHF, which inhibits the function of TS which in turn depletes intracellular thymidylate resulting in suppression of DNA synthesis. Again, 5'-formylTHF, the precursor of 5, 10-methyleneTHF potentiates the cytotoxic effects of 5-FU by stabilizing the inhibition of 5, 10-methyleneTHF-TS-5FdUMP complex. Studies have shown that MTHFR polymorphisms that increase the concentration of 5, 10-methyleneTHF may result in elevation of the cytotoxic levels of 5FU. Conversely, MTHFR polymorphisms that increase of accumulation of 5,10-methyleneTHF may alter the cytotoxic effects of Methotrexate (MTX), another common chemotherapeutic agent. MTX inhibits dihydrofolate reductase, decreases intracellular 5,10-methyleneTHF levels for the synthesis of thymidylate and inhibits the synthesis of purine. Accumulation of dihydrofolate can inhibit the activities of TS and other enzyme. These chemotherapeutic agents affect DNA synthesis and down-regulate cell replication. Experimental studies on breast cancer cell lines, mice models and breast cancer patients have shown that MTHFR polymorphisms influence the concentration and cellular pathway of folate metabolism altering their chemosensitivity.

Some of the common chemotherapy medicines used currently for breast cancer treatment are abraxane, adriamycin, carboplatin, cytoxan, daunorubicin, doxil, ellence, fluorouracil also called 5-fluorouracil or 5-FU, gemzar, halaven, ixempra, methotrexate, mitomycin, navelbine, taxol, taxotere, thiopeta,
vincristine, and xeloda\textsuperscript{62, 63}. These chemotherapy medicines are, however, commonly used in combinations. The type of chemotherapy regimen recommended to a patient is dependent upon the patient's general health, menopausal status and clinical characteristics like the stage at diagnosis, hormone receptor status, HER2 status, and lymph node status. Some of the current standard chemotherapy regimens\textsuperscript{62, 63} used for breast cancer treatment include:

- Adriamycin and Taxotere (AT)
- Adriamycin and Cytoxan with or without Taxol or Taxotere (AC± T)
- Cytoxan, Methotrexate, and Fluorouracil (CMF)
- Cytoxan, Ellence, and Fluorouracil (CEF)
- Fluorouracil, Adriamycin, and Cytoxan (FAC)
- Cytoxan, Adriamycin, and Fluorouracil (CAF)
- Taxotere, Adriamycin, and Cytoxan (TAC)
- Gemzar, Ellemce, and Taxol (GET)

Studies have not yet investigated the effect of various commonly used chemotherapy medicines on the association of BHMT, MTR, MTRR, SHMT, and FOLR1 and breast cancer survival. In this study, therefore, it was investigated if chemotherapy modified the associations of these genes with breast cancer survival.

**ER status:**

ER is a group of receptors that are activated by estrogen hormone\textsuperscript{64}. The main function of ER is to serve as a DNA binding transcription factor that
regulates gene expression. There were two reasons for investigating the effects of ER status on the association between genes in the folate metabolizing pathway and breast cancer survival. First, DNA methylation has been suggested in the literature to be associated with silencing of ER expression. The enzymes in the folate metabolizing pathway may influence ER status. Secondly, it is often hypothesized that tumorigenesis is triggered by binding of estrogen to the ER which stimulates proliferation of mammary cells decreasing the rate of cell division, DNA replication and mutation. Either way, normal cell cycle and DNA repair mechanisms are disrupted resulting in inhibition of apoptosis and tumor formation. ER-status is, therefore, a widely used prognostic marker and indicator of breast cancer treatment. ER expression is an important marker for deciding whether or not a patient receives endocrine therapy (for example tamoxifen) as part of breast cancer treatment regimen.

In the Long Island Breast cancer study a significant elevation of mortality rates was observed due primarily to breast cancer among C677T (MTHFR) allele carriers compared to non-carriers. Additionally, ER/PR status modified the association of C677T polymorphism with all-cause mortality after breast cancer treatment. Martin et al also observed the effects of ER status on the association between both the MTHFR polymorphisms and breast cancer-specific mortality. In their sample, women with ER negative tumors had worse survival.
COVARIATES OF INTEREST

1. Folate intake:

Folate is essential for DNA methylation and nucleotide synthesis. Folate deficiency could, therefore, lead to abnormal DNA methylation\(^\text{10}\). Epidemiological studies have shown that deficiency in folate could lead to breast\(^\text{11-17}\), colorectal\(^\text{71-73}\), and endometrial\(^\text{74}\) cancer and the effect of the deficiency is modified by \textit{MTHFR} polymorphisms\(^\text{11-15,71-74}\). Recent studies on colorectal cancer have also shown that while folate in general is anti-neoplastic before the establishment of tumor foci, after the tumor is established it accelerates tumor proliferation\(^\text{75,76}\). While high consumption of folate reduces risk for colorectal cancer by 30-40% compared to lowest consumption of folate, timing of intake is crucial because the literature suggests that progression of premalignant to cancerous lesions is accelerated with high folate intake\(^\text{77,78}\). Since folate provides one carbon unit in the thymidylate synthase mediated conversion of dUMP to dTMP, its depletion could result in the accumulation of intracellular uracil resulting in the mis-incorporation of uracil into DNA and increase in DNA damage\(^\text{79,80}\).

2. Stage at diagnosis:

Stage at diagnosis is another widely used prognostic marker that is an indicators for breast cancer treatment type\(^\text{81}\) and for survival. Shrubsole et al\(^\text{12}\) in their study on \textit{MTHFR} polymorphisms observed a non-significant elevation of risk (HR= 1.8; 95% CI: 0.79-4.14) for mortality due to breast cancer among TT genotype compared to carriers of CC genotype who were diagnosed at a later
stage (stage III-IV) and the risk was almost three fold higher (HR= 2.97; 95% CI: 1.10-7.98) among those that survived longer than two years following initial diagnosis. Results from both the Long Island Breast Cancer Study\textsuperscript{26} and the Shanghai Breast Cancer Study\textsuperscript{20}, indicated that C677T genotype significantly elevated the risk for breast cancer-specific mortality among patients diagnosed with late stage (stage III-IV) disease. The study by Martin and colleagues did not stratify on disease stage\textsuperscript{18}.

3. Treatments type:

Besides chemotherapy, tamoxifen therapy was a covariate of interest in this study. It is estimated that approximately 70% of breast cancer cases are ER positive\textsuperscript{70} and ER status is a clinical deciding factor for prescription of tamoxifen and aromatase inhibitors\textsuperscript{82}. ER bound to estradiol and tamoxifen induces different conformations in the receptor accounting for variation in functional activity. The drug tamoxifen, thus acts as an antagonist in breast cells and hence ideal for breast cancer treatment\textsuperscript{70}. As mentioned already, DNA methylation may influence the expression of ER\textsuperscript{66-69}.

4. Demographic factors:

Demographic factors like age, race/ethnicity, body mass index (BMI), reproductive factors like age at menarche, age at menopause, parity, etc., often used as surrogates for exposures to endogenous hormones, and exogenous hormones like OC and HRT use have been widely recognized as factors that influence survival following diagnosis and treatment of breast cancer\textsuperscript{83-85}. Although effect modification of the association between breast cancer risk and
MTHFR by menopausal status has been investigated, the findings of the existing literature have been non-significant and inconsistent. Studies have not yet investigated the influence of menopausal status on breast cancer survival among women carrying various other polymorphisms in folate metabolizing pathway and genes responsible for folate intake. These covariates are, therefore, potential confounders in the investigation of the association between genes in the folate metabolizing pathway and breast cancer-specific survival.

5. Other nutritional and lifestyle factors:

Biochemical studies have shown that B-complex vitamins are involved in the pathogenesis of breast cancer in several pathways. Deficiencies in vitamin B₆ (pyridoxine), B₁₂ (cyanocobalamin), niacin causes DNA breakage and oxidative lesion. Both vitamin B₆ and B₁₂ are directly involved in the folate metabolizing pathway (Figure 1). For instance, SHMT1 enzyme is dependent on vitamin B₆ for purine synthesis necessary for conversion of homocysteine to methionene and also catalyzes the conversion from 5-methyltetrahydrofolate to 5-10-methyltetrahydrofolate. Vitamin B₁₂ is directly involved in the methylation process via SAM synthesis.

Studies have also shown the interaction of alcohol consumption and survival following diagnosis of breast cancer. Although the risk for early mortality due to breast cancer among women with 40 g/day alcohol consumption is elevated (HR= 2.00; 95% CI: 1.14-3.49) with low folate intake (200μg/day), high levels (400μg/day) of folate intake seems to have a protective effect against mortality (HR= 0.77; 95% CI: 0.33-1.80) due to breast cancer.
Dietary and lifestyle factors along with prognostic factors serve as potential confounders of the association between folate intake and survival following breast cancer diagnosis and treatment. Polymorphisms in the folate metabolizing pathway alter the concentration and cellular pathway of folate metabolism disrupting the sensitivity of chemotherapeutic drugs used in breast cancer treatment.

6. Comorbidities:

Studies have investigated the associations of the genes in folate metabolizing pathway on the etiology of colon cancer\textsuperscript{21,22,76,93}. It was therefore, intriguing to explore the percentage of the deaths due to cancer other than breast cancer in the HEAL sample can be attributable to colon cancer and if they serve as confounds. It was therefore, important to consider deaths due to other cancers that might confound the association between the genes in the folate metabolizing pathway and breast cancer survival.

Studies have also shown that higher concentrations of homocysteine, an amino acid involved in SAM synthesis (see figure 1), elevate the risk of cardiovascular disease (CVD)\textsuperscript{94,95}. In presence of vitamin B\textsubscript{12}, \textit{MTR}, \textit{BHMT}, \textit{MTRR}, and \textit{MTHFR} catalyzes the conversion of homocysteine to methionine by remethylation. Under normal physiological conditions homocysteine is excreted by the kidney in moderate amounts\textsuperscript{96} but polymorphisms of genes involved in homocysteine metabolisms could increase the plasma concentration of homocysteine\textsuperscript{97,98}, which is thought to damage the vascular epithelium and thereby increase the risk for CVD\textsuperscript{99,100}. Studies have also shown that vitamin B\textsubscript{12}
and folate are protective against CVD development\textsuperscript{101}. In the literature it is suggested that patients with vascular diseases and atherosclerosis have lower levels of folate and vitamin B12 concentrations\textsuperscript{102} and vitamin B\textsubscript{12} deficiency may elevate the homocysteine concentration and contribute plaque formation in the carotid artery\textsuperscript{103}. In other words, vitamin B\textsubscript{12} deficiency is an important risk factor for both hyperhomocysteinemia and CVD\textsuperscript{95}. 
CHAPTER III
METHODS AND STATISTICAL PLAN

INTRODUCTION

This chapter presents the methods used for this dissertation. The chapter begins with a brief overview of the study population and the samples used in the study. It is followed by a list of study hypotheses, a description of the Cox Proportional Hazard method, primary dependent and independent variables and censorings used in the survival analysis. The section is followed by a list of the operational hypotheses and a description of the main effects and the interactions investigated in this dissertation project. There is a section on power analysis section and concludes with brief description of the genotyping process.

STUDY POPULATION

The Health, Eating, Activity, and Lifestyle (HEAL) Study is a population based, multicenter, multiethnic prospective cohort study in which 1,183 women with incident primary BC were enrolled to investigate the associations of weight, physical activity, diet, sex hormones, mammographic density, and other established risk factors with BC survival and quality of life. Women were recruited through three Surveillance, Epidemiology, and End Results (SEER) registries: at the Fred Hutchinson Cancer Research Center in Seattle, WA, the University of New Mexico and the University of Southern California. Details of the study aims
design, and recruitment procedures have been published previously \(^{104-106}\).

Briefly, in the New Mexico site a total of 615 women, age 18 years or older, diagnosed with *in situ* to stage IIIA BC between July 1996 and March 1999 and living in Bernalillo, Sante Fe, Sandoval, Valencia, or Taos Counties were recruited. In the Seattle site, 202 women, between the age of 40 and 64 years, diagnosed with *in situ* to stage IIIA BC between September 1997 and September 1998 and living in King, Pierce, or Snohomish Counties were recruited. In Los Angeles (LA) site, 366 African American women with stage 0 to IIIA primary BC who either participated in the LA portion of the Women’s Contraceptive and Reproductive Experiences Study, which is a case-control study of invasive BC, or in a parallel case control study of *in situ* BC. Other eligibility criteria for the LA site included were born in the US, capable of speaking English, between the ages of 35 to 64 years and diagnosed with BC between May 1995 and May 1998. Participants completed an in-person interview at baseline which was within their first year following diagnosis (an average of 7.5 months after diagnosis) and at 24 months after the baseline visit which was within their third year of diagnosis (an average of 31 months after diagnosis). During both time points biological specimens were collected. Data were collected again at 5 and 10 years. Person years of follow-up are calculated from the date of diagnosis to 12/31/2007 as the end point. For investigation of the SNPs associated with *MTHFR* gene samples from all three sites were used. A diagrammatic representation of the cohort is presented in Chapter 4.
To investigate the association of the SNPs corresponding to \textit{BHMT}, \textit{MTR}, \textit{MTRR}, \textit{SHMT1}, and \textit{FOLR1} genes and breast cancer mortality only the samples from New Mexico site was available. Of the 615 women enrolled at New Mexico, DNA was collected for 406 women at baseline and 409 at 24 months follow-up. Because DNA was collected twice, some women who did not provide samples at baseline provided them at follow-up. Thus, the total number of women with DNA available for this study was 462. Three hundred and sixty-eight had provided DNA during baseline as well as at 24 months follow-up interview, 36 at baseline only and 58 at follow-up only. During the preparation phase, all 416 samples were inventoried and sample selection was based on availability with first preference given to baseline DNA and, if a sample did not have enough DNA from baseline, the sample from the 24 months follow-up was used. For 12 samples, there were insufficient amounts of DNA from both baseline and 24 months follow-up to perform genotyping and one sample was compromised for an unknown reason and could not be genotyped for all five SNPs. A total of 446 biological samples were used for the current investigation of which 137 (31\%) were collected during baseline and 309 (69\%) were from 24 months follow-up interview. Figure 8 provides a diagrammatic representation of the cohort.
Enrollment at baseline: 615

Biological specimens collected:
Baseline: 406
Follow-up: 409

Archived DNA available:
Baseline and follow up: 368
Baseline only: 36
Follow-up only: 58
Total: 462

Insufficient DNA available: 12
Unable to genotype: 4
Total DNA lost: 16

Current study:
DNA (baseline): 137
DNA (24 months follow-up): 309
Total DNA used: 446

Mortality (11.46 years follow-up):
All causes: 67
Breast cancer-specific: 22

Figure 8 Cohort definition for the New Mexico site of the Health, Eating, Activity, and Lifestyle (HEAL) Study and the number of events during 11.46 years of follow-up
HYPOTHESIS

Below are the study hypotheses:

1. **Main effect:**
   
   Genes in one carbon metabolizing pathway (*MTHFR, BHMT, SHMT1, MTR, MTRR, and FOLR1*) are responsible for modulating key enzymes in folate metabolisms. Several case control studies\textsuperscript{11-17} have investigated the association of *MTHFR* polymorphisms and breast cancer risk and breast cancer survival\textsuperscript{12 18 19} but only one cohort study has so far investigated the association of the other genes in the folate metabolizing pathway and breast cancer survival. In this study, therefore, we test the following hypothesis:

   **Women who carry at least one minor allele (dominant model) of the candidate* TAG SNP polymorphisms experience an increase in mortality; specifically breast cancer mortality within the first 11.46 years of follow-up compared to women who carry two major alleles adjusting for relevant covariates.**

   For example, carriers of the variant allele of C677T, a polymorphism of *MTHFR* gene, are reported reduced of breast cancer specific mortality rates\textsuperscript{19}.

2. **Gene-Gene interaction:**

   Previous studies have not investigated the association of genes within the folate metabolizing pathway and breast cancer-specific survival. In this study, therefore, we test the following hypothesis:
Presence of other genes in the same pathway increases the all-cause, and breast cancer mortality rates among women who carried at least one minor allele of the candidate/TAG (BHMT, SHMT1, MTR, MTRR, and FOLR1) polymorphisms.

3. Gene-Environment interaction:

a) The current literature suggests possible biological associations between chemotherapy and breast cancer survival \(^{60, 61}\). Genes in the folate metabolizing pathway may alter the levels of 5-10 methyleneTHF which may increase the levels of THF leading to increase in the cytotoxic levels of folate depleting chemotherapeutic agents, 5FU and MTX\(^{60, 61}\). In this study, therefore, the following hypothesis was tested:

Women who received chemotherapy and carried at least one minor allele (dominant model) of the candidate/TAG (BHMT, SHMT1, MTR, MTRR, and FOLR1) polymorphisms have increased rates of all-cause, and breast cancer mortality compared to women who did not receive chemotherapy and carried two major alleles.

b) The assessment of ER status influences the type of treatment a women receives which may influence her survival\(^{28, 82, 107}\). DNA methylation has been suggested to be associated with silencing of ER expression\(^{66-69}\). Results from the Long Island Breast Cancer Study \(^{19}\) suggested a significant interaction between MTHFR (C677T) and ER status for breast cancer survival. In this study, therefore, we test the following hypothesis:
Women diagnosed with ER negative tumors and carrying at least one minor allele (dominant model) of the candidate TAG (BHMT, SHMT1, MTR, MTRR, and FOLR1) polymorphisms have increased rates of all-cause, and breast cancer mortality compared to women with ER positive tumor who carried two major alleles.

STATISTICAL PLAN

To test the study hypotheses, Cox proportional-hazards (PH) analysis (SAS 9.2, PROC PHREG) was used to estimate all cause and breast cancer-specific mortality for main effect and interaction with chemotherapy and ER status.

The formula for the Cox PH model:

\[
\frac{\sum_{i=1}^{p} \beta_i x_i}{h(t, X) = h_0(t) e^{\sum_{i=1}^{p} \beta_i x_i}}
\]

Here \( h_0(t) \) is the baseline hazard at time \( t = 0 \) and \( X = (X_1, X_2, \ldots X_p) \) are predictor variables that are independent of time. Here \( h_0(t) \) is unspecified. Since the hazard function involves \( t \) and an exponential expression involving the \( X \) without \( t \), it makes fitting the model non-negative. It is, therefore, robust non-parametric, non-linear and non-negative model\(^{108}\). Parameters of the Cox model were obtained by maximizing the likelihood function which is the joint probability of obtaining the data that is observed on the study subjects as a function of the unknown parameter’s (\( \beta \)) of the model under consideration\(^{108}\).
Variables:

Multivariate Cox proportional-hazards models were created for each of the six SNPs and potentially important nutritional, behavioral and/or demographic covariates to investigate gene-environment interaction. Following are the primary dependent and independent variables.

Primary dependent variables or outcomes of interest:

1. Overall survival/all-cause mortality
2. Breast cancer-specific mortality

Primary independent variables:

1. Genotypes of interest: Genes (MTHFR (polymorphisms: C677T and A1298C), BHMT, MTR, MTRR, SHMT1, FOLR1 and RFC1) are coded as homozygous dominant (AA) = 0, heterozygous (AB) = 1 and homozygous recessive (BB) = 2.
2. Chemotherapy use: Although various chemotherapeutic agents were received either alone or in various combinations owing to sample size considerations (see chapter 4), chemotherapy was coded as a binary variable where no chemotherapy= 0 and chemotherapy= 1.
3. Estrogen receptor status: Estrogen receptor (ER) status of the tumor was re-coded into ER-positive= 0 and ER- negative = 1.

Covariates:

Covariates of interest were adjusted for all models based on biological plausibility, and frequency distribution. Relevant covariates were subjected to
correlation and leverage analysis prior to being incorporated into the models. Covariates of interest based include:

Potential confounders were adjusted in all models based on biological plausibility and frequency distribution. Relevant covariates were subjected to correlation and leverage analysis prior to being incorporated into the models. Covariates of interest included:

1. Age at diagnosis (continuous): Based on the eligibility criteria, all subjects in the study were between ages 29 to 87 years at the time of breast cancer diagnosis.

2. Ethnicity: Ethnicity was coded as Non-Hispanic White = 0 (reference), and Hispanic= 1.

3. Body Mass Index: The continuous variable BMI was created based on body weight and height. For the purpose of analysis it was recoded into four categories based on WHO recommendations: BMI <18.5 (underweight), BMI between 18.5 and 24.9 (normal), BMI between 25.0 and 29.9 (overweight) and BMI >= 30.0 (obese). The variable underweight was considered as the reference category.

4. Menopausal status: Menopausal status was coded as a binary variable with pre-menopausal= 0 and post-menopausal= 1.

5. Stage at diagnosis: Based on eligibility criteria only cases with in situ to stage IIIA of breast cancer were included in the study. The stage of disease at diagnosis, was based on SEER defined,
Summary Stage 1977, composed of three sub-groups: in situ, localized, and regional.

6. Treatment: Treatment was categorized as: no surgery = 0, surgery only = 1, surgery and radiation = 2, surgery and chemotherapy = 3 and surgery, radiation and chemotherapy = 4.

7. Tamoxifen use: Tamoxifen use was coded as a binary variable; previous or present treated was coded as 1 and no treatment as 0.

8. Radiation therapy: Radiation therapy was coded as a binary variable; previous radiation was coded as 1 and no treatment as 0.

9. Oral contraceptive (OC) use: OC use was coded as a binary variable. If a subject reported receiving OC she was coded as 1 else 0.

10. Hormonal therapy (HRT): HRT was coded as a binary variable. If a subject reported receiving HRT she was coded as 1 else 0.

11. Age at menarche: Age at menarche was used as a continuous variable.

12. Age at menopause: Age at menopause was used as a continuous variable.

13. Parity: Parity was coded as 0 for nulliparous subjects, and 1 for subjects with one or more live births.

14. Family history of breast cancer: Family history of breast cancer was coded such that if a subject reported that she has no family history of breast cancer diagnosis she was coded as 0, she was coded as
1 if any 1st degree relative(s) were diagnosed with breast cancer and 2 if any 2nd degree relative(s) and no 1st degree relative were diagnosed with breast cancer.

15. Smoking status: Smoking is also coded as a binary variable. If a subject reported smoking the past or during the time of the interview (ever/current smoker) she was coded as 1. If she reported never smoking in her life she is coded as 0.

16. Alcohol intake: Like total folate intake variable, amounts of alcohol intake reported during 24 months interview was considered in the analysis. It was reported in mg/day of intake. Total alcohol intake was considered as continuous variable. For the purpose of analysis it was also re-coded in tertiles. All outliers were excluded from the analysis.

17. Folate intake: Since Los Angeles site did not collect food intake data at baseline, data collected during 24 months interview was used in the analysis for investigating the two MTHFR polymorphisms. Given that there was strong correlation between baseline and 24-months total folate intake for the New Mexico and Seattle data justified the used of the 24 months food frequency data. But for investigating the other five genes, where only New Mexico samples were used baseline food frequency data were used. It was reported in µg/day of intake. Total folate intake was considered as a continuous variable. For the purpose of analysis it
was also re-coded into tertiles where the lowest tertile was considered the reference category. All outliers (> 2000 µg/day) were excluded from the analysis.

18. B complex vitamin intake (B₁, B₂, B₃, B₆ and B₁₂): Like folate, the data for the rest of the B-complex vitamins, namely: niacin (B₁), riboflavin (B₂), thiamin (B₃), vitamin-B₆, and vitamin-B₁₂, were also used from the 24 months interview. It was reported in µg/day of intake. All vitamin B variables were considered as continuous variables. For the purpose of analysis it was also recoded into tertiles.

19. Total energy intake: Information on total energy intake was also used from 24 months interview data. It was reported in Kcal/day. It was used as a continuous variable. It was also recoded into tertiles for the purpose of analysis. All outliers (> 13000 µg/day) were excluded from the analysis.

Censoring:

A censored observation is said to occur when there is incomplete information for survival time. Following are the reasons for censoring to occur in the study.

Primary outcomes:

For all-cause mortality:

1. End of follow-up: the subject does not experience death due to any cause before the study ends
2. Lost to follow-up: the subject withdraws from the study because of reasons not related to the study outcome (person moving to another location and not leaving any forwarding address, busy, not-interested in participating in the study, etc.)

For breast cancer specific mortality:

1. End of follow-up: the subject does not experience death due to breast cancer before the study ends
2. Lost to follow-up: the subject withdraws from the study because of reasons not related to the study (person moving to another location and not leaving any forwarding address, busy, not-interested in participating in the study, etc.)

Although four different kinds of censoring are possible: right truncation, left truncation, right censoring and left censoring, this study focuses only on right censoring because the central idea behind survival analysis is to follow subjects over time and observe at which point in time the event of interest occurs, which in this study is all-cause death or breast cancer death. Thus, when an observation is right censored it means that the information is incomplete because an event was not experienced during the study time period or the individual was lost to follow-up or withdrew from the study.

OPERATIONAL HYPOTHESIS

Main effect:

Equation for the multivariable model:

\[ h(t,X) = h_0(t) \exp \left[ \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots \beta_p X_p \right] \]
Where $X_1$ is the gene under investigation (Example: $MTHFR$ (C677T); 0 = CC (wild type) and 1 = CT/TT (variant)) and $X_2$ through $X_p$ are covariates of interest that are adjusted for in the model. The coefficients $\beta_1$ through $\beta_p$ describes the main effects of the primary independent variables and covariates of interest.

**Interaction:**

Multivariable interaction model:

$$h(t,X) = h_0(t) \exp \left[ \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 \ast X_2) + \beta_4 X_4 \right. \left. + \ldots + \beta_p X_p \right]$$

Where $X_1$ is the gene under investigation (Example: $MTHFR$ (C677T); 0 = CC (wild type) and 1 = CT/TT (variant)) and $X_2$ is a primary independent variable (Example: chemotherapy, 0 = no and 1 = yes). $X_4$ through $X_p$ are covariates of interest that are adjusted for in the model. The coefficients $\beta_1$ and $\beta_2$ describes the main effects of the gene and the primary independent variables respectively, $\beta_3$ describes the increase or decrease in hazard ratio due to interactions between the gene and the primary independent variable.

Multivariate interaction model testing dose or types of a treatment:

$$h(t,X) = h_0(t) \exp \left[ b X + c_1 Z_1 + \ldots + c_{k-1} Z_{k-1} + d_1 (X \ast Z_1) + \ldots + d_{k-1} (X \ast Z_{k-1}) \right]$$

Where $X_1$ is the gene under investigation (Example: $MTHFR$ (C677T); 0 = CC (wild type) and 1 = CT/TT (variant)), $Z_1 \ldots Z_{k-1}$ are types or levels of the primary...
independent variable (Example: types of chemotherapy). The coefficients $b$ and $c_1 \ldots c_{k-1}$ describe the main effects of the gene and types or levels of the primary independent variable respectively, and $d_1 \ldots d_{k-1}$ describe increase or decrease in hazard ratio due to the interactions of the gene and types or levels of the primary independent variable.

Gene-gene interaction was investigated based on allele and haplotype frequencies (PROC ALLELE and PROC HAPLOTYPE) of the genotype subgroups. Haplotype analysis was focused only on the two polymorphs of \textit{MTHFR} gene. Covariates of interest were adjusted for all models based on biological plausibility, and frequency distribution. Relevant covariates were subjected to correlation and leverage analysis prior to being incorporated into the models.

**POWER CALCULATIONS**

The estimated minor allele frequencies of one carbon metabolizing genes in the general population are reported in table 3. The selection of the SNPs in each of the five genes is based on the following criteria:

- Estimated allele frequency of the SNP is greater than 10 % and less than 40 %
- The studies that helped estimate the allele frequencies had a sample size greater than 1000
- Allele frequencies used for power calculation (reported in table 3) were estimated based on studies which had samples that spread across the globe.
Power equations:

If \( f \) is the genotype frequency of interest and \( g=1-f \), and \( D \) is the number of deaths after 10 years of follow-up (in this study), power to detect a hazard ratio \( \Delta \) if we use a one-sided test is calculated using the following equations. Here it is assumed that there are no ties in survival times and the power of the log-rank test depends on the sample size which in turn depends the number of deaths. \(^{109}\) Let \( Z_{1-\alpha} \) be the percentile corresponding to \( \alpha \% \) probabilities, \( P [ Z > Z_{1-\alpha} ] = \alpha \) and \( P [ Z < Z_{1-\alpha} ] = 1-\alpha \), where \( \alpha = 1-\alpha \) or \( 1-\beta \), then from Hsieh et al we have,

\[
D = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{f g [\log \Delta]^2}
\]

\[\text{(1)}\]

\[
=> D \times f g \times [\log \Delta]^2 = (Z_{1-\alpha} + Z_{1-\beta})^2
\]

\[\text{(2)}\]

For a two-sided test, where \( \alpha = 5\% \) and \( 1-\beta = 80\% \), \( Z_{1-\alpha} = 1.96 \) then

\[
P [ Z < Z_{1-\alpha} ] = P [ Z < 1.96 ] = 0.975
\]

and

\[
P [ Z > Z_{1-\alpha} ] = P [ Z > 1.96 ] = 0.025
\]

For a one sided test, \( Z_{1-\alpha} = 1.64 \) then

\[
P [ Z < Z_{1-\alpha} ] = P [ Z < 1.64 ] = 0.
\]

and

\[
P [ Z > Z_{1-\alpha} ] = P [ Z > 1.64 ] = 0.20
\]
Genotype frequencies for the dominant and recessive models:

In genetics a trait is considered dominant, when it expresses a phenotypic effect with only one copy of the gene and it is considered recessive if a trait is expressed only if two copies of the gene are inherited. To illustrate, if we consider the gene BHMT (A742G), the population allele frequency for the major allele (A) and minor alleles (G) are 0.28 and 0.72 (1- 0.28= 0.72) respectively. If Hardy Weinberg equilibrium ($p^2 + 2pq + q^2= 1$) is applied:

\[
\begin{align*}
p^2 & = GG = (0.72)^2 = 0.52 \\
2pq & = GA = 2 \times 0.72 \times 0.28 = 0.40 \\
q^2 & = AA = (0.28)^2 = 0.08 
\end{align*}
\]

For the recessive model,

\[
\begin{align*}
GG + GA & = 0.52 + 0.40 = 0.92 \\
AA & = 0.08 
\end{align*}
\]

For the dominant model,

\[
\begin{align*}
GG & = 0.52 \\
GA + AA & = 0.40 + 0.08 = 0.48 
\end{align*}
\]

The following table shows the values of f and g corresponding to dominant and recessive model for all five SNPs in this study
Table 2: Values of \( f \) and \( g \) for dominant and recessive models

<table>
<thead>
<tr>
<th>Gene</th>
<th>Allele freq</th>
<th>( f )</th>
<th>( g )</th>
<th>( f )</th>
<th>( g )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHMT (A742G)</td>
<td>0.28</td>
<td>GG/GA</td>
<td>0.92</td>
<td>AA</td>
<td>0.08</td>
</tr>
<tr>
<td>MTR (A2756G)</td>
<td>0.19</td>
<td>AA/AG</td>
<td>0.96</td>
<td>GG</td>
<td>0.04</td>
</tr>
<tr>
<td>MTRR (G66A)</td>
<td>0.35</td>
<td>AA/AG</td>
<td>0.88</td>
<td>GG</td>
<td>0.12</td>
</tr>
<tr>
<td>SHMT1 (G1420A)</td>
<td>0.26</td>
<td>AA/AG</td>
<td>0.93</td>
<td>GG</td>
<td>0.07</td>
</tr>
<tr>
<td>FOLR1 (G606A)</td>
<td>0.32</td>
<td>GG/GA</td>
<td>0.90</td>
<td>AA</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Power for overall and stratified (chemotherapy Vs no chemotherapy) models:

For a one tailed test, by inserting the values corresponding to \( BHMT \) gene in equation 2 \((Z_{1-\alpha} = 1.64)\), to detect a hazard ratio of 2.00 for a total of 107 deaths out of the 772 subjects enrolled in the study, we have:

\[
Z_{1-\beta} = \sqrt{Dfg} \times \log\Delta - Z_{1-\alpha} \\
= \sqrt{107 \times 0.48 \times 0.52} \times \log(2) - 1.64 \\
= 1.94
\]

So \( 1-\beta = 0.974 \) which is 97.4% power.

For a two-sided test:

\[
= \sqrt{107 \times 0.48 \times 0.52} \times \log(2) - 1.96 \\
= 1.62
\]

So \( 1-\beta = 0.948 \) which is 94.8% power.
Applying the same equation the results for all five SNPs are reported in table 3. Microsoft Excel 2010 was used to analyze powers to detect a hazard ratio of 2.00, 1.75 and 1.50 for both two and one tailed tests. Power for stratified models (chemotherapy v. no chemotherapy) was also calculated for both two and one tailed tests and is shown in table 4.

**Table 3: Power (two and one-tailed) to test association between genotype and breast cancer mortality for the main effects of the dominant models using all three HEAL sites (N= 772)**

<table>
<thead>
<tr>
<th>Gene</th>
<th>RECESSIVE MODEL</th>
<th></th>
<th></th>
<th></th>
<th>DOMINANT MODEL</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Allele freq</td>
<td>HR 2.00</td>
<td>HR 1.75</td>
<td>HR 1.50</td>
<td>HR 2.00</td>
<td>HR 1.75</td>
<td>HR 1.50</td>
<td></td>
</tr>
<tr>
<td>BHMT (A742G)</td>
<td>0.28</td>
<td>1.00</td>
<td>0.99</td>
<td>0.95</td>
<td>0.95</td>
<td>0.82</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>MTR (A2756G)</td>
<td>0.19</td>
<td>0.9</td>
<td>0.21</td>
<td>0.13</td>
<td>0.93</td>
<td>0.79</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>MTRR (G66A)</td>
<td>0.35</td>
<td>0.64</td>
<td>0.47</td>
<td>0.28</td>
<td>0.94</td>
<td>0.82</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>SHMT1 (G1420A)</td>
<td>0.26</td>
<td>0.45</td>
<td>0.32</td>
<td>0.19</td>
<td>0.95</td>
<td>0.82</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>FOLR1 (G606A)</td>
<td>0.32</td>
<td>0.58</td>
<td>0.41</td>
<td>0.24</td>
<td>0.95</td>
<td>0.82</td>
<td>0.55</td>
<td></td>
</tr>
</tbody>
</table>

**Two-tailed test**

<table>
<thead>
<tr>
<th>Genes</th>
<th>Allele freq</th>
<th>HR 2.00</th>
<th>HR 1.75</th>
<th>HR 1.50</th>
<th>HR 2.00</th>
<th>HR 1.75</th>
<th>HR 1.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHMT (A742G)</td>
<td>0.28</td>
<td>1.00</td>
<td>1.00</td>
<td>0.98</td>
<td>0.97</td>
<td>0.89</td>
<td>0.68</td>
</tr>
<tr>
<td>MTR (A2756G)</td>
<td>0.19</td>
<td>0.41</td>
<td>0.30</td>
<td>0.21</td>
<td>0.96</td>
<td>0.87</td>
<td>0.64</td>
</tr>
<tr>
<td>MTRR (G66A)</td>
<td>0.35</td>
<td>0.76</td>
<td>0.59</td>
<td>0.39</td>
<td>0.97</td>
<td>0.89</td>
<td>0.67</td>
</tr>
<tr>
<td>SHMT1 (G1420A)</td>
<td>0.26</td>
<td>0.58</td>
<td>0.44</td>
<td>0.28</td>
<td>0.97</td>
<td>0.89</td>
<td>0.67</td>
</tr>
<tr>
<td>FOLR1 (G606A)</td>
<td>0.32</td>
<td>0.69</td>
<td>0.54</td>
<td>0.35</td>
<td>0.97</td>
<td>0.89</td>
<td>0.67</td>
</tr>
</tbody>
</table>

**One-tailed test**
Table 4: Power (two and one-tailed) to test association between genotype and breast cancer mortality for the dominant models stratified by use of chemotherapy using all three HEAL sites

<table>
<thead>
<tr>
<th>Gene</th>
<th>Allele freq</th>
<th>RECESSIVE MODEL</th>
<th>DOMINANT MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Chemotherapy (N=315)</td>
<td>No Chemotherapy (N=457)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HR 2.00 HR 1.75 HR 1.50</td>
<td>HR 2.00 HR 1.75 HR 1.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two-tailed test</td>
<td></td>
</tr>
<tr>
<td>BHMT (A742G)</td>
<td>28</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>MTR (A2756G)</td>
<td>19</td>
<td>0.21</td>
<td>0.15</td>
</tr>
<tr>
<td>MTRR (G66A)</td>
<td>35</td>
<td>0.49</td>
<td>0.35</td>
</tr>
<tr>
<td>SHMT1 (G1420A)</td>
<td>26</td>
<td>0.33</td>
<td>0.23</td>
</tr>
<tr>
<td>FOLR1 (G606A)</td>
<td>32</td>
<td>0.43</td>
<td>0.30</td>
</tr>
</tbody>
</table>

One-tailed test

| BHMT (A742G) | 28 | 1.00 | 0.99 | 0.91 | 0.96 | 0.87 | 0.64 | 0.91 | 0.78 | 0.54 | 0.64 | 0.49 | 0.32 |
| MTR (A2756G) | 19 | 0.32 | 0.24 | 0.17 | 0.19 | 0.15 | 0.12 | 0.89 | 0.74 | 0.51 | 0.60 | 0.46 | 0.29 |
| MTRR (G66A) | 35 | 0.62 | 0.47 | 0.31 | 0.37 | 0.28 | 0.19 | 0.90 | 0.77 | 0.53 | 0.63 | 0.48 | 0.31 |
| SHMT1 (G1420A) | 26 | 0.45 | 0.34 | 0.23 | 0.27 | 0.21 | 0.15 | 0.91 | 0.78 | 0.54 | 0.63 | 0.48 | 0.32 |
| FOLR1 (G606A) | 32 | 0.56 | 0.42 | 0.28 | 0.33 | 0.25 | 0.17 | 0.91 | 0.78 | 0.54 | 0.64 | 0.49 | 0.32 |
If the calculations are restricted to samples from the New Mexico site only, power for the main effect and stratified models (chemotherapy v. no chemotherapy) for both two and one-tailed tests are shown in Tables 5 and 6 respectively.

**Table 5: Power (two and one-tailed) to test association between genotype and breast cancer mortality for the main effects of the dominant models using samples from New Mexico site (N= 446)**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Allele freq</th>
<th>RECESSIVE MODEL</th>
<th></th>
<th></th>
<th></th>
<th>DOMINANT MODEL</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HR 2.00</td>
<td>HR 1.75</td>
<td>HR 1.50</td>
<td>HR 2.00</td>
<td>HR 1.75</td>
<td>HR 1.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BHMT</td>
<td>0.28</td>
<td>0.97</td>
<td>0.87</td>
<td>0.61</td>
<td>0.60</td>
<td>0.43</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A742G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTR</td>
<td>0.19</td>
<td>0.14</td>
<td>0.10</td>
<td>0.07</td>
<td>0.56</td>
<td>0.40</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A2756G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTRR</td>
<td>0.35</td>
<td>0.30</td>
<td>0.21</td>
<td>0.13</td>
<td>0.59</td>
<td>0.42</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G66A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHMT1</td>
<td>0.26</td>
<td>0.20</td>
<td>0.15</td>
<td>0.09</td>
<td>0.59</td>
<td>0.43</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G1420A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOLR1</td>
<td>0.32</td>
<td>0.26</td>
<td>0.19</td>
<td>0.12</td>
<td>0.59</td>
<td>0.43</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G606A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Two-tailed test**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Allele freq</th>
<th>RECESSIVE MODEL</th>
<th></th>
<th></th>
<th></th>
<th>DOMINANT MODEL</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BHMT</td>
<td>0.28</td>
<td>0.99</td>
<td>0.92</td>
<td>0.72</td>
<td>0.72</td>
<td>0.56</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A742G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTR</td>
<td>0.19</td>
<td>0.22</td>
<td>0.17</td>
<td>0.13</td>
<td>0.68</td>
<td>0.53</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A2756G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTRR</td>
<td>0.35</td>
<td>0.42</td>
<td>0.32</td>
<td>0.21</td>
<td>0.71</td>
<td>0.55</td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G66A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHMT1</td>
<td>0.26</td>
<td>0.30</td>
<td>0.23</td>
<td>0.16</td>
<td>0.72</td>
<td>0.56</td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G1420A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOLR1</td>
<td>0.32</td>
<td>0.38</td>
<td>0.29</td>
<td>0.19</td>
<td>0.72</td>
<td>0.56</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G606A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6: Power (two and one-tailed) to test association between genotype and breast cancer mortality for the dominant models stratified by use of chemotherapy using samples from New Mexico site (N= 446)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Allele freq</th>
<th>DOMINANT MODEL</th>
<th>RECESSIVE MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Chemotherapy</td>
<td>No Chemotherapy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(N=118)</td>
<td>(N=288)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HR=2.00 HR=1.75 HR=1.50</td>
<td>HR=2.00 HR=1.75 HR=1.50</td>
</tr>
<tr>
<td>BHMT (A742G)</td>
<td>0.28</td>
<td>0.83 0.65 0.39</td>
<td>0.69 0.51 0.30</td>
</tr>
<tr>
<td>MTR (A2756G)</td>
<td>0.19</td>
<td>0.09 0.08 0.06</td>
<td>0.80 0.06 0.05</td>
</tr>
<tr>
<td>MTRR (G66A)</td>
<td>0.35</td>
<td>0.19 0.14 0.09</td>
<td>0.15 0.11 0.08</td>
</tr>
<tr>
<td>SHMT1 (G1420A)</td>
<td>0.26</td>
<td>0.14 0.10 0.07</td>
<td>0.11 0.09 0.06</td>
</tr>
<tr>
<td>FOLR1 (G606A)</td>
<td>0.32</td>
<td>0.17 0.13 0.09</td>
<td>0.14 0.10 0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One-tailed test</td>
<td></td>
</tr>
<tr>
<td>BHMT (A742G)</td>
<td>0.28</td>
<td>0.89 0.76 0.53</td>
<td>0.79 0.63 0.42</td>
</tr>
<tr>
<td>MTR (A2756G)</td>
<td>0.19</td>
<td>0.17 0.14 0.11</td>
<td>0.14 0.12 0.09</td>
</tr>
<tr>
<td>MTRR (G66A)</td>
<td>0.35</td>
<td>0.29 0.23 0.16</td>
<td>0.24 0.19 0.14</td>
</tr>
<tr>
<td>SHMT1 (G1420A)</td>
<td>0.26</td>
<td>0.22 0.17 0.13</td>
<td>0.18 0.15 0.11</td>
</tr>
<tr>
<td>FOLR1 (G606A)</td>
<td>0.32</td>
<td>0.27 0.21 0.15</td>
<td>0.22 0.17 0.13</td>
</tr>
</tbody>
</table>
Limitations of the power analysis:

For the purpose of this study, however, a one tailed test is used for the following reasons:

- Due to limited current literature on the association between the genes in the folate metabolizing pathway and breast cancer survival, the direction of the associations could not be determined \textit{a priori}.
- A one sided test has greater power over a two-sided non-directional test.

Power analysis was based on the entire HEAL cohort based on data from the preliminary investigation of the two polymorphs corresponding to the \textit{MTHFR} gene (Chapter IV). The preliminary study included biological specimens from all three sites collected during baseline interview. Samples diagnosed with \textit{in situ} tumors were excluded from the study (refer to Chapter IV for exclusion criteria). But for studying the association between the SNPs corresponding to \textit{BHMT, MTR, MTRR, SHMT1, and FOLR1} and breast cancer mortality only the samples from New Mexico site were available. Blood specimens collected during both baseline and 24 months follow-up interviews were used for genotyping of these SNPs.

\textbf{GENOTYPING}

For the New Mexico site, 30mL of blood was collected from women who consented to provide biological specimens at baseline and 24 month follow-up visits. The blood samples were processed and stored at $-70$ to $-80^\circ\text{C}$ within 3 hours of collection. DNA was isolated from stored buffy coat using Purogene (Gentra Systems, Inc.) extraction kits. A StepOnePlus real-time polymerase
chain reaction (RT-PCR) system (Applied Biosystems by Life Technologies Corporations) was used to genotype the five SNPs under investigation. The RT-PCR was factory calibrated to generate accurate outputs with 99.7% confidence\textsuperscript{110}. TaqMan SNP genotyping assays were used consisting of two PCR primers specific to the locus under investigation and two oligonucleotide probes specific to the allele of interest. Each probe has fluorescent dye, FAM and VIC at the 5' end and a non-fluorescent quencher with a minor groove binder (MGB) at the 3' end\textsuperscript{111}. The major allele of a SNP was identified by the FAM dye while the minor allele was identified by the VIC dye. A reaction cocktail of 10 μL of 2X master mix, 0.5 μL 40X assay, and 4.5 μL of de-ionized water was used with required amounts of DNA-TE mixture to initiate each RT-PCR reaction. 

During each RT-PCR cycle, the probe, which was hydrolyzed to the allele of interest, emitted a fluorescent signal when it cleaved to the 5' exonuclease action of the AmpliTaq Gold DNA Polymerase. The primers amplified the locus of interest in the DNA template and the dye specific to the alleles reported the presence or absence of the allele in the DNA sample\textsuperscript{112,113}. The presence of two probes, each specific to an allele of the SNP under investigation and labeled with two dyes, enabled two alleles to be detected in a single tube\textsuperscript{113}. Appendix III provides the details of the genotyping procedure.

After genotyping was completed, results were tested for the Hardy Weinberg Equilibrium ($p^2 + 2pq + q^2 = 1$). Table 7 shows the calculations for the Hardy Weinberg Equilibrium in this study. There was significant difference
between the observed and the expected frequencies as is reflected in the chi square p values.

**Table 7: Calculations for Hardy Weinberg Equilibrium for the SNPs investigated in this study**

<table>
<thead>
<tr>
<th>Genes</th>
<th>Obs AA</th>
<th>Obs AB</th>
<th>Obs BB</th>
<th>Exp AA</th>
<th>Exp AB</th>
<th>Exp BB</th>
<th>p</th>
<th>q</th>
<th>Chi-sq</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHMT</td>
<td>204</td>
<td>198</td>
<td>44</td>
<td>205.8</td>
<td>194.3</td>
<td>45.8</td>
<td>0.679</td>
<td>0.321</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>MTR</td>
<td>280</td>
<td>138</td>
<td>25</td>
<td>274.9</td>
<td>148.1</td>
<td>19.9</td>
<td>0.788</td>
<td>0.212</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>MTRR</td>
<td>104</td>
<td>216</td>
<td>121</td>
<td>101.9</td>
<td>220.2</td>
<td>118.9</td>
<td>0.481</td>
<td>0.519</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>SHMT</td>
<td>211</td>
<td>194</td>
<td>38</td>
<td>214.1</td>
<td>187.7</td>
<td>41.1</td>
<td>0.695</td>
<td>0.305</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>FOLR</td>
<td>210</td>
<td>187</td>
<td>48</td>
<td>206.9</td>
<td>193.0</td>
<td>44.9</td>
<td>0.682</td>
<td>0.318</td>
<td>0.81</td>
<td></td>
</tr>
</tbody>
</table>

Obs AA: observed homozygous wild-type; Obs AB: observed heterozygous; Obs BB: observed homozygous variant; Exp AA: expected homozygous wild-type; Exp AB: expected heterozygous; Exp BB: expected homozygous variant.
CHAPTER IV
PRELIMINARY STUDY:
MTHFR POLYMORPHISMS AND BREAST CANCER SURVIVAL

INTRODUCTION

The preliminary investigation reports results from the analyses of the associations between the C677T and A1298C MTHFR polymorphisms with breast cancer-specific survival and how chemotherapy modifies the effect of the polymorphisms on survival using data from the HEAL Study.

MATERIALS AND METHODS

Statistical methods

Cox Proportional Hazard analysis was applied to compare the effect of MTHFR polymorphisms on women who received chemotherapy and those who did not receive chemotherapy on three outcomes: all-cause mortality and breast cancer-specific mortality. Analyses adjusted for relevant covariates including stage at diagnosis, ER-status, menopausal status, site, race, BMI, tamoxifen use, and family history of BC. Because race and site were highly correlated, site was included in the final model. Menopausal status, BMI, tamoxifen use and family history of BC did not influence the effect estimates and were not included in the final model. Univariate models for the two MTHFR polymorphisms, chemotherapy use, stage at diagnosis, ER-status and site were constructed for
the three outcomes. Since stage at diagnosis, ER-status and chemotherapy are highly correlated, only chemotherapy was used in the final model. Multivariate models tested for interactions between the variant alleles and chemotherapy. Chemotherapy was stratified into two categories based on folate-dependence and used in multivariate models. Diploftype and compound homozygous and heterozygous models were created to test if the effect of combined genotype modifies chemotherapy use on the three outcomes.

RESULTS

Below is a diagrammatic representation of the cohort. Also shown in the figure is the number of all-cause and breast cancer specific deaths stratified by chemotherapy. Only subjects diagnosed with invasive tumors with biological specimens were used in the analysis.
Figure 9 Health, Eating, Activity, and Lifestyle (HEAL) cohort definition and number of events during 10 years of follow-up stratified by chemotherapy

Figure 9 shows the distribution of relevant demographic, genotype and clinical characteristics for the cohort stratified by treatment with chemotherapy (yes, no). Fifty-four percent of the participants were non-Hispanic white, 31% African American and 13% Hispanic. Sixty-two percent were post-menopausal, and 35% were overweight or obese. Forty-eight percent had the T allele (CT/TT) for C677T and 45% had the C allele (AC/CC) for A1298C. Those who received
Chemotherapy tended to be African American, younger, premenopausal women with more advanced disease and ER negative tumors at diagnosis.

**Table 8: Demographic and medical characteristics of breast cancer survivors (N= 772) in the Health, Eating, Activity, and Lifestyle study**

<table>
<thead>
<tr>
<th>Descriptive Characteristics</th>
<th>No Chemotherapy N=457 %</th>
<th>Chemotherapy N=315 %</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C677T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>229 (50)</td>
<td>175 (56)</td>
<td>0.137</td>
</tr>
<tr>
<td>CT/TT</td>
<td>228 (50)</td>
<td>140 (44)</td>
<td></td>
</tr>
<tr>
<td>A1298C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>245 (54)</td>
<td>181 (57)</td>
<td>0.291</td>
</tr>
<tr>
<td>AC/CC</td>
<td>212 (46)</td>
<td>134 (53)</td>
<td></td>
</tr>
<tr>
<td>Age at Diagnosis^2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58.3 (11.41)</td>
<td>49.2 (8.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Mass Index (Kg/m^2)</td>
<td></td>
<td></td>
<td>0.333</td>
</tr>
<tr>
<td>Normal (&lt; 25.0)</td>
<td>283 (62)</td>
<td>213 (68)</td>
<td></td>
</tr>
<tr>
<td>Over weight (25.0-29.9)</td>
<td>99 (22)</td>
<td>57 (18)</td>
<td></td>
</tr>
<tr>
<td>Obese (&gt;=30)</td>
<td>70 (15)</td>
<td>45 (14)</td>
<td></td>
</tr>
<tr>
<td>Study site</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>New Mexico</td>
<td>288 (63)</td>
<td>118 (37)</td>
<td></td>
</tr>
<tr>
<td>Seattle</td>
<td>64 (14)</td>
<td>63 (20)</td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>105 (23)</td>
<td>134 (43)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hispanic</td>
<td>61 (13)</td>
<td>37 (12)</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>281 (61)</td>
<td>135 (.43)</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>105 (23)</td>
<td>134 (43)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>10 (3)</td>
<td>9 (2)</td>
<td></td>
</tr>
<tr>
<td>Menopausal Status</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Pre-menopause</td>
<td>113 (26)</td>
<td>158 (54)</td>
<td></td>
</tr>
<tr>
<td>Post-menopause</td>
<td>317 (74)</td>
<td>133 (46)</td>
<td></td>
</tr>
<tr>
<td>Breast cancer stage</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Localized</td>
<td>401 (88)</td>
<td>134 (43)</td>
<td></td>
</tr>
<tr>
<td>Regional</td>
<td>54 (12)</td>
<td>181 (57)</td>
<td></td>
</tr>
<tr>
<td>ER Status</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Positive</td>
<td>345 (87)</td>
<td>168 (57)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>53 (13)</td>
<td>125 (43)</td>
<td></td>
</tr>
<tr>
<td>PR Status</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Positive</td>
<td>267 (73)</td>
<td>142 (53)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>97 (27)</td>
<td>128 (47)</td>
<td></td>
</tr>
</tbody>
</table>

^Cell size may vary because of missing values. Only participants in the HEAL study diagnosed with invasive tumor and with available DNA from the 24-months interview were used in the analysis; ^2 mean ± SD
Table 9 shows that carriers of the variant alleles of both C667T and A1298C polymorphisms, had lower breast cancer mortality compared to the homozygous wild-type (C677T: HR= 0.76, 95 % CI: 0.51, 1.11; A1298C: HR= 0.71, 95 % CI: 0.48, 1.05).

Table 9: Crude Hazard Ratios for any cause mortality and breast cancer-specific mortality for study variables (11.46 years)

<table>
<thead>
<tr>
<th>Variables</th>
<th>ALL CAUSE MORTALITY</th>
<th>BREAST CANCER-SPECIFIC MORTALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaths</td>
<td>Hazard Ratio 95% CI</td>
</tr>
<tr>
<td>C677T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>96 55</td>
<td>1.00</td>
</tr>
<tr>
<td>CT/TT</td>
<td>79 45</td>
<td>0.89 0.66-1.20</td>
</tr>
<tr>
<td>A1298C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>105 60</td>
<td>1.00</td>
</tr>
<tr>
<td>AC/CC</td>
<td>70 40</td>
<td>0.79 0.59-1.07</td>
</tr>
<tr>
<td>Chemo therapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>90 51</td>
<td>1.66 &lt;0.01</td>
</tr>
<tr>
<td>Yes</td>
<td>1.23-2.23</td>
<td>2.31-5.25</td>
</tr>
</tbody>
</table>

*Information on ER-status and stage at diagnosis was not available for all subjects
Table 10 Prognosis: stratified by chemotherapy and genotype showing breast cancer mortality

<table>
<thead>
<tr>
<th>Chemo Therapy</th>
<th>Genotype</th>
<th>Deaths N</th>
<th>%</th>
<th>Deaths Ratio</th>
<th>95% CI</th>
<th>P-value</th>
<th>Deaths N</th>
<th>%</th>
<th>Deaths Ratio</th>
<th>95% CI</th>
<th>P-value</th>
<th>Deaths N</th>
<th>%</th>
<th>Deaths Ratio</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C677T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>11</td>
<td>12</td>
<td>Reference</td>
<td>2.24</td>
<td>0.06</td>
<td>16</td>
<td>18</td>
<td>Reference</td>
<td>2.26</td>
<td>0.05</td>
<td>16</td>
<td>15</td>
<td>Reference</td>
<td>1.58</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>CT/TT</td>
<td>16</td>
<td>18</td>
<td></td>
<td>0.97-5.17</td>
<td></td>
<td>16</td>
<td>16</td>
<td></td>
<td>0.98-5.22</td>
<td></td>
<td>17</td>
<td>16</td>
<td></td>
<td>0.75-3.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>40</td>
<td>45</td>
<td>Reference</td>
<td>0.75</td>
<td>0.32</td>
<td>22</td>
<td>25</td>
<td>Reference</td>
<td>0.70</td>
<td>0.19</td>
<td>47</td>
<td>44</td>
<td>Reference</td>
<td>0.78</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>CT/TT</td>
<td>22</td>
<td>25</td>
<td></td>
<td>0.43-1.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.42-1.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.47-1.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1298C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>18</td>
<td>20</td>
<td>Reference</td>
<td>0.65</td>
<td>0.31</td>
<td>9</td>
<td>10</td>
<td>Reference</td>
<td>0.64</td>
<td>0.29</td>
<td>18</td>
<td>19</td>
<td>Reference</td>
<td>0.76</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>AC/CC</td>
<td>39</td>
<td>44</td>
<td>Reference</td>
<td>0.28-1.49</td>
<td></td>
<td>26</td>
<td>27</td>
<td>Reference</td>
<td>0.26-1.47</td>
<td></td>
<td>44</td>
<td>45</td>
<td>Reference</td>
<td>0.37-1.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>23</td>
<td>26</td>
<td></td>
<td>0.82</td>
<td>0.46</td>
<td>23</td>
<td>26</td>
<td></td>
<td>0.82</td>
<td>0.43</td>
<td>28</td>
<td>29</td>
<td></td>
<td>0.86</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>AC/CC</td>
<td>26</td>
<td>27</td>
<td></td>
<td>0.49-1.39</td>
<td></td>
<td>27</td>
<td>28</td>
<td></td>
<td>0.50-1.34</td>
<td></td>
<td>29</td>
<td>30</td>
<td></td>
<td>0.53-1.38</td>
<td></td>
</tr>
</tbody>
</table>

*Adjusted for age at diagnosis, and race. Stage at diagnosis and ER status are not adjusted for in the model due to strong correlation with chemotherapy.
Table 11: Interaction of genotype with chemotherapy on breast cancer mortality

<table>
<thead>
<tr>
<th>Treatment</th>
<th>C677T</th>
<th>A1298C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FIVE YEARS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Deaths</td>
<td>Deaths</td>
</tr>
<tr>
<td></td>
<td>N %</td>
<td>95% CI</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>P for the interaction</td>
<td>0.13</td>
<td>0.72</td>
</tr>
</tbody>
</table>

| **SEVEN YEARS** | | |
| No | Deaths | Deaths | Deaths | Deaths | Deaths | Deaths |
| | N % | 95% CI | N % | 95% CI | N % | 95% CI | N % | 95% CI |
| Chemotherapy | 40 | 45 | 16 | 18 | 1.73 | 0.17 | 18 | 20 | Reference | 9 | 10 | 0.67 |
| Chemotherapy | 40 | 45 | 16 | 18 | 1.73 | 0.17 | 18 | 20 | Reference | 9 | 10 | 0.67 |
| P for the interaction | 0.08 | 0.71 |

| **TEN YEARS** | | |
| No | Deaths | Deaths | Deaths | Deaths | Deaths | Deaths |
| | N % | 95% CI | N % | 95% CI | N % | 95% CI | N % | 95% CI |
| Chemotherapy | 40 | 45 | 16 | 18 | 1.73 | 0.17 | 18 | 20 | Reference | 9 | 10 | 0.67 |
| Chemotherapy | 40 | 45 | 16 | 18 | 1.73 | 0.17 | 18 | 20 | Reference | 9 | 10 | 0.67 |
| P for the interaction | 0.37 | 0.95 |

*Adjusted for age at diagnosis and race.
Table 10 shows results for analyses of the interaction of the variant alleles in C677T and A1298T with chemotherapy by duration of follow-up, adjusting for age, study site, race, stage of disease and ER status. In a multiplicative model, in the absence of interaction the expected HR for those with the minor allele (CT/T or AC/CC) who received chemotherapy should be approximately equal to the HR for those with minor allele who did not received chemotherapy multiplied by the HR for the homozygous wild-type (CC or AA) group who received chemotherapy. The results, however, suggest antagonistic, multiplicative interaction between the T allele of C677T and chemotherapy because the observed HR for their joint presence is substantially less than expected. For example, for 7 years of follow-up, the expected result should be 1.18 (HR for CT/T not receiving chemotherapy) x 3.33 (HR for CC receiving chemotherapy) = 3.93 (HR_{exp} for CT/T receiving chemotherapy). The observed HR, however, is 2.51, or 30% less than the expected result, and the formal test of interaction also approaches statistical significance (p = 0.09). The results in Table 11 also suggest that the interaction between the C677T polymorphism and chemotherapy is stronger during the first 5 to 7 years and attenuates somewhat thereafter. This is reflected in the p-values for interaction as well as in comparison of the expected to observed HRs for the joint effect group; for example the expected HR for CT/T and chemotherapy at 10 years is 4.86 compared to the observed value of 3.25. We also computed an interaction model for C677T with chemotherapy without adjustment for stage of disease and ER status. While the absolute magnitudes of the HRs were larger across all
strata, the relative size of the interaction was similar (the observed HR was approximately 45% less than the expected value). In contrast to C677T, there was no evidence whatsoever for interaction between A1298C and chemotherapy at 5, 7 or 10 years of follow-up.

Because stage of disease and ER status are criteria used by clinical oncologists to assign breast cancer patients to receive chemotherapy rather than hormonal therapy, the possibility exists for residual confounding of the interaction between C677T and chemotherapy by these prognostic factors. As a result, we also explored the interactions of the MTHFR polymorphisms with stage of disease (local versus regional) and ER (negative versus positive) independent of chemotherapy. None of the interactions were statistically significant or approached significance (p’s > 0.40; Hazard Ratios not shown) for at 5, 7 or 10 years of follow-up. Once again, there was no evidence for interaction between the A1298C polymorphism and either stage of disease or ER status on breast cancer mortality.

The Kaplan Meier graphs shown in Figures 10 and 11 provide additional evidence for an interaction between C677T and chemotherapy. The difference between chemotherapy and no chemotherapy groups is visibly greater in the women with the CC genotype than in those with CT/TT genotypes, or T allele.

To further investigate the effect of chemotherapy on breast cancer-specific, chemotherapy was stratified based on folate dependence (folate depleting chemotherapy (170 (54%)) versus non-folate depleting chemotherapy (145(46%)) and adjusted for relevant covariates. No significant change in hazard
ratios were observed (C677T: HR= 1.75, 95% CI: 0.79, 3.84; A1298C: HR=0.63, 95% CI: 0.28, 1.42).

**Figure 10**: Kaplan-Meier curve for ten years breast cancer-specific mortality by chemotherapy for carriers of CC genotype (C677T polymorphism) showing number of months from diagnosis to breast cancer death using data from Health, Eating, Activity, and Life-style study (N= 772)

![Kaplan-Meier curve for CC genotype](image)

**Figure 11**: Kaplan-Meier curve for ten years breast cancer-specific mortality by chemotherapy for carriers of CT/TT genotype (C677T polymorphism) showing number of months from diagnosis to breast cancer death using data from Health, Eating, Activity, and Life-style study (N= 772)

![Kaplan-Meier curve for CT/TT genotype](image)
Figure 12: Kaplan-Meier curve for ten years breast cancer-specific mortality for carriers of A1298C polymorphism showing number of months from diagnosis to breast cancer death using data from Health, Eating, Activity, and Life-style study (N= 772)

DISCUSSION

The results of this investigation suggest that carriers of the variant alleles of both the C677T and A1298C MTHFR polymorphisms may be protected against breast cancer mortality. Our results additionally suggest that the C677T polymorphism may modify breast cancer survival in women receiving chemotherapy. While formal tests of statistical significance for both main effects and interactions of these MTHFR polymorphisms with chemotherapy were not significant due the small sample size (n = 772), the effect sizes observed are potentially, clinically meaningful and should be replicated in larger cohort studies or clinical trials. These polymorphisms have been established to modify folate metabolism, which is important for DNA methylation and repair, and have been previously reported to be associated with breast cancer survival as well as breast
cancer risk\textsuperscript{11-17}. Because common chemotherapeutic agents used in breast cancer, such as MTX and 5FU, also impact folate metabolism, we hypothesized that these polymorphisms might also modify the efficacy of these drugs.

Biochemical studies have shown that besides being an anti-metabolite of folate, two metabolites of 5-FU, 5-fluoro-2′-deoxyuridine-5′-triphosphate and 5′-flourouridine-5′-triphosphate when incorporated into DNA and RNA respectively could interfere with the level of DNA stability and RNA functioning\textsuperscript{60}. Again, 5′-formylTHF, the precursor of 5, 10-methyleneTHF potentiates the cytotoxic effects of 5-FU by stabilizing the inhibition of 5, 10-methyleneTHF-TS- 5FdUMP complex. Studies have shown that the \textit{MTHFR} polymorphisms that increases the concentration of 5, 10 -methyleneTHF may result in elevation the cytotoxic levels of 5FU\textsuperscript{60}. Conversely, \textit{MTHFR} polymorphisms that increases accumulation of 5,10 -metheleneTHF is shown to alter the cytotoxic effects of MTX\textsuperscript{61}. MTX inhibits dihydrofolate reductase, decreases intracellular 5,10 -methyleneTHF levels for the synthesis of thymidylate and inhibits the synthesis of purine\textsuperscript{61}. In short, these chemotherapeutic agents affect DNA synthesis and intimately down-regulate cell replication. In fact, experimental studies investigating the relationship between \textit{MTHFR} polymorphisms and commonly used chemotherapeutic agents on breast cancer cell lines\textsuperscript{23,24}, mice models\textsuperscript{24} and breast cancer patients\textsuperscript{20,25-27} have shown that \textit{MTHFR} polymorphisms influences the concentration and cellular pathway of folate metabolism altering their chemosensitivity.
The results suggest that the C677T polymorphism may modify the effects of chemotherapeutic treatment. This interpretation is supported by the following arguments. The presence of the T allele for C677T was not associated stage at diagnosis or ER status, and the interaction with chemotherapy was not meaningfully diminished by adjusting for these covariates of chemotherapy. Moreover, it is difficult to see how the T allele might modify the associations of these prognostic variables with subsequent survival because they are fixed characteristics established at the time of diagnosis. Adjusting for these variables diminished the absolute values of the HRs, but did not affect the relative strength of interaction, indicating confounding between chemotherapy and these prognostic characteristics. Thus, the better than expected survival of women receiving chemotherapy who have the T allele may reflect a modification of the effectiveness of chemotherapeutic treatment. Once again, the plausibility of this interpretation is supported by the fact that the women receiving chemotherapy were predominantly taking methotrexate and/or 5FU, which are known to affect folate metabolism.

The data suggests that there is no correlation between the C677T and chemotherapy (Pearson correlation coefficient; \( r = -0.003 \)), so it may support the notion that the variant allele of the C677T alter the metabolism of the chemotherapeutic drug, reducing the probability of adverse events or may reduce the side effects of the drug, thus contributing to decrease in breast cancer mortality among carriers of the T allele compared to those not carrying the T allele. The direct association of A1298C, on the other hand, suggests protection
against breast cancer mortality, even after adjusting for established prognostic markers of breast cancer. A1298C may influence breast cancer survival through a different pathway from C677T.

To test the association between the two polymorphisms and breast cancer-specific mortality we performed diplotype analysis. The results of the diplotype analysis or compound genotype analysis suggest although recombination between the variant alleles of the two polymorphisms was rare (CTAC; n=130, TTAC; n=3, TTCC; n=0) (figure 12). There was strong linkage disequilibrium between the two polymorphisms (Lewontin’s D' = 0.94) suggesting that the two polymorphisms may be associated with breast cancer survival in two different ancestral pathways. The fact that there was no correlation (0.001) between them further supports the assumption.

![Figure 12 Patterns of recombination showing change in C677T and A1298C polymorphisms](image)

Although no correlation among the polymorphisms and ER-status, and stage at diagnosis was observed, there were strong correlations among both
covariates and chemotherapy because they are significant prognostic markers of breast cancer and are important deciding factors in chemotherapeutic regimen. Both ER-status and stage at diagnosis may, therefore, be mediators of the relationship between genotype and chemotherapy. The notion is further supported by the results of the univariate analyses which show approximately three times higher breast cancer-specific mortality rates for all three variables. Both covariates were, therefore, not adjusted in our multivariate models. The effects of ER-status and stage at diagnosis were investigated in this study. Unlike the results of Long Island Breast Cancer Study\textsuperscript{19} and Shanghai Breast Cancer Study\textsuperscript{20} which suggested that ER-status modified the effect of C677T and breast cancer survival and C677T significantly increased the survival of women diagnosed with late stage (stage III-IV) tumor respectively, no such interactions were observed in this study.
CHAPTER V
RESULTS

INTRODUCTION

The current chapter describes the findings of the dissertation. In all the tables in this chapter the two polymorphisms of \textit{MTHFR} are also listed along with the five genes investigated in this dissertation project. The purpose of having the two \textit{MTHFR} polymorphisms was to compare the findings of this analysis based on samples from the New Mexico site with findings based on an earlier analysis using all available samples from the three HEAL study sites (see Chapter IV), and (b) to test for gene-gene interaction between the \textit{MTHFR} polymorphisms and breast cancer-specific mortality.

DESCRIPTIVE STATISTICS

Out of the 446 women in the New Mexico site, a total of 67 (15\%) were deceased as of 12/31/2007 at 11.46 years; 22 (5\%) deaths were due to breast cancer for women diagnosed with invasive disease. More than three-fourths (78\%) of the participants were non-Hispanic white, 22\% were Hispanic and 3\% belonged to the other category. The majority of the participants were normal (43\%; BMI: 18.5-24.9) or over-weight (31\%; BMI: 25.0-29.9), 8\% were below normal (BMI: <18.5) and 18\% were obese (BMI: \textit{>=}30). Sixty-eight percent of the sample was post-menopausal and 32 \% were pre-menopausal. More than half of the participants were diagnosed with a localized tumor (62\%), 18\% were
diagnosed with regional tumor and 20% with *in situ* breast disease. Table 12 describes the demographic and medical characteristics of the participants (N= 446) in the New Mexico site and table 13 provides the frequency distribution of the genotypes of the breast cancer survivors.

**Table 12: Demographic and medical characteristics of breast cancer survivors for the New Mexico site in the Health, Eating, Activity, and Lifestyle Study (n= 446)**

<table>
<thead>
<tr>
<th>Descriptive Characteristics</th>
<th>Breast Cancer Deaths</th>
<th>Total N</th>
<th>%/SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at Diagnosis (yrs)</strong></td>
<td>22</td>
<td>57.6</td>
<td>12.1</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>97</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>346</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td><strong>Body Mass Index (Kg/m²)</strong></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below normal (&lt;18.5)</td>
<td>37</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Normal (18.5-24.9)</td>
<td>190</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Over weight (25.0-29.9)</td>
<td>137</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Obese (≥30)</td>
<td>78</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td><strong>Menopausal Status</strong></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-menopause</td>
<td>136</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Post-menopause</td>
<td>294</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td><strong>Stage at Diagnosis</strong></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In situ</td>
<td>89</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Localized</td>
<td>276</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>Regional</td>
<td>81</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td><strong>Estrogen Receptor Status</strong></td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>257</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>56</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td><strong>Progesterone Receptors Status</strong></td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>213</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>89</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td><strong>Chemotherapy</strong></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>334</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>112</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td><strong>Tamoxifen use</strong></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>242</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>204</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td><strong>Oral Contraceptive Use</strong></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>160</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>286</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Hormone replacement Therapy&lt;sup&gt;7&lt;/sup&gt;</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>186</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>259</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>Family History of Breast Cancer&lt;sup&gt;8&lt;/sup&gt;</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>197</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; degree</td>
<td>121</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; degree</td>
<td>93</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Current Smoker&lt;sup&gt;9&lt;/sup&gt;</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>201</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>48</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Total Folate Intake (µg)&lt;sup&gt;10&lt;/sup&gt;</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>316.2</td>
<td>126.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamins (mg/d)&lt;sup&gt;10&lt;/sup&gt;</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total B1 (Thiamin)</td>
<td>1.1</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Total B2 (Riboflavin)</td>
<td>1.6</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>Total B3 (Niacin)</td>
<td>14.9</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Total B6 (pyridoxine, pyridoxal, &amp; pyridoxamine)</td>
<td>1.4</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>Total B12 (Cobalamin)</td>
<td>4.5</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>Alcohol (gm/day)&lt;sup&gt;10&lt;/sup&gt;</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>9.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy Intake (kcal)&lt;sup&gt;10&lt;/sup&gt;</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1401.3</td>
<td>611</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup>Percentages calculation excludes missing values; <sup>2</sup> mean ± SD; <sup>3</sup> other in race; <sup>4</sup> missing values in BMI; <sup>5</sup> missing values in ER-status: 133; <sup>6</sup> missing values in PR-status: 144; <sup>7</sup> missing values in OC and HRT use: 1; <sup>8</sup> missing values in family history of breast cancer: 35; <sup>9</sup> missing values among current smoker: 197; <sup>10</sup> missing values in energy intake, vitamins, folate and alcohol intake: 14.

Table 13: Genotypes of breast cancer survivors for the New Mexico site in the Health, Eating, Activity, and Lifestyle Study (n= 446)

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Breast Cancer Deaths</th>
<th>Total N</th>
<th>%</th>
<th>P value HWE&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHMT (N= 446; rs3733890)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A742G</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>22</td>
<td>204</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>198</td>
<td>44</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>242</td>
<td>25</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>AG/GG</td>
<td></td>
<td>163</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>MTR (N= 443; rs1805087)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2756G</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>21</td>
<td>280</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>138</td>
<td>25</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>163</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG/GG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Sample Size</td>
<td>rsID</td>
<td>Allele Frequency</td>
<td>Genotype Distribution</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
<td>------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>MTRR (N= 441; rs1801394)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G66A</td>
<td>21</td>
<td></td>
<td></td>
<td>GG 104 / AG 216 / AA 307</td>
</tr>
<tr>
<td>SHMT1 (N= 443; rs1979277)</td>
<td></td>
<td></td>
<td></td>
<td>GG 211 / AG 194 / AA 38</td>
</tr>
<tr>
<td>C1420T/ G1420A</td>
<td>22</td>
<td></td>
<td></td>
<td>GG 210 / AG 197 / AA 48</td>
</tr>
<tr>
<td>FOLR1 (N= 445; rs3796517)</td>
<td></td>
<td></td>
<td></td>
<td>GG 165 / CT 184 / TT 48</td>
</tr>
<tr>
<td>G606A</td>
<td>22</td>
<td></td>
<td></td>
<td>CC 165 / CT 184 / TT 48</td>
</tr>
<tr>
<td>MTHFR (N= 397; rs1801133)</td>
<td></td>
<td></td>
<td></td>
<td>CC 197 / AC 165 / CC 35</td>
</tr>
<tr>
<td>C677T</td>
<td>20</td>
<td></td>
<td></td>
<td>CC 165 / CT 184 / TT 48</td>
</tr>
<tr>
<td>MTHFR (N= 397; rs1801131)</td>
<td></td>
<td></td>
<td></td>
<td>AA 197 / AC 165 / CC 35</td>
</tr>
<tr>
<td>A1298C</td>
<td>20</td>
<td></td>
<td></td>
<td>AA 197 / AC 165 / CC 35</td>
</tr>
</tbody>
</table>

*Abbreviation: HWE: Hardy Weinberg Equilibrium.*
AIM 1: MAIN EFFECTS

The associations between the genes in the folate metabolizing pathway and all-cause and breast cancer-specific mortality are summarized in table 14. All models reported in table 14 were adjusted for age. Adjustments for the potential confounders, race/ethnicity, menopausal status, BMI, stage at diagnosis, treatment types, ER status, folate, B-vitamin intake, and alcohol intake did not change the point estimates by more than 10%.
Table 14 Age-adjusted hazard ratios\(^1\) and 95% confidence intervals for the associations of polymorphisms in the folate metabolizing pathway and all-cause mortality and breast cancer-specific mortality using data from the New Mexico site of the Health, Eating, Activity, and Lifestyle study (N= 446).

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>(\text{ALL-CAUSE MORTALITY}^{\text{Deaths N, Censored N %, HR (95% CI), P value}})</th>
<th>(\text{BREAST CANCER-SPECIFIC MORTALITY}^{\text{Deaths N, Censored N %, HR (95% CI), P value}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHMT (N= 446) A742G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>30 174 85 1.00</td>
<td>8 196 96 1.00</td>
</tr>
<tr>
<td>AG</td>
<td>26 172 87 0.98 0.6-1.7 0.93</td>
<td>12 189 94 1.52 0.6-3.7 0.36</td>
</tr>
<tr>
<td>GG</td>
<td>11 33 75 1.61 0.8-3.2 0.18</td>
<td>2 42 95 1.17 0.3-5.5 0.84</td>
</tr>
<tr>
<td>AG/ GG</td>
<td>37 205 85 1.10 0.7-1.8 0.89</td>
<td>14 228 94 1.45 0.6-3.5 0.39</td>
</tr>
<tr>
<td>MTR (N= 443) A2756G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>42 238 85 1.00</td>
<td>14 266 95 1.00</td>
</tr>
<tr>
<td>AG</td>
<td>21 117 85 1.19 0.7-2.0 0.52</td>
<td>6 132 96 0.82 0.3-2.1 0.82</td>
</tr>
<tr>
<td>GG</td>
<td>3 22 88 0.79 0.3-2.6 0.69</td>
<td>1 24 96 0.82 0.1-6.2 0.85</td>
</tr>
<tr>
<td>AG/ GG</td>
<td>24 139 85 1.12 0.7-1.9 0.67</td>
<td>7 156 96 0.82 0.3-2.0 0.67</td>
</tr>
<tr>
<td>MTRR (N= 441) G66A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>27 94 78 1.00</td>
<td>10 111 92 1.00</td>
</tr>
<tr>
<td>GA</td>
<td>28 188 87 0.52 0.3-0.9 0.01</td>
<td>7 209 97 0.38 0.2-1.0 0.05</td>
</tr>
<tr>
<td>AA</td>
<td>11 93 89 0.47 0.2-0.9 0.03</td>
<td>4 100 96 0.42 0.1-1.4 0.15</td>
</tr>
<tr>
<td>GA/ AA</td>
<td>39 281 88 0.51 0.3-0.8 &lt;0.01</td>
<td>11 309 97 0.40 0.2-0.9 0.04</td>
</tr>
<tr>
<td>SHMT1 (N= 443) G1420A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>35 176 83 1.00</td>
<td>10 201 95 1.00</td>
</tr>
<tr>
<td>AG</td>
<td>29 165 85 0.93 0.6-1.5 0.78</td>
<td>11 183 94 1.17 0.5-2.8 0.71</td>
</tr>
<tr>
<td>AA</td>
<td>3 35 92 0.57 0.2-1.9 0.35</td>
<td>1 37 97 0.51 0.1-3.9 0.52</td>
</tr>
<tr>
<td>GA/ AA</td>
<td>32 200 86 0.87 0.5-1.4 0.57</td>
<td>12 220 95 1.06 0.5-2.5 0.89</td>
</tr>
</tbody>
</table>

\(^1\) Adjusting for clinical covariates stage at diagnosis, chemotherapy use, ER-status did not affect point estimates by more than 10%
Table 14 continued: Age-adjusted hazard ratios\textsuperscript{1} and 95% confidence intervals for the associations of polymorphisms in folate metabolizing pathway and all-cause mortality and breast cancer-specific mortality using data from the New Mexico site of the Health, Eating, Activity, and Lifestyle study (N= 446).

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>ALL-CAUSE MORTALITY</th>
<th>BREAST CANCER-SPECIFIC MORTALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaths N</td>
<td>Censored N</td>
</tr>
<tr>
<td>FOLR1(N= 446)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G606A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>27</td>
<td>183</td>
</tr>
<tr>
<td>AG</td>
<td>32</td>
<td>155</td>
</tr>
<tr>
<td>AA</td>
<td>8</td>
<td>40</td>
</tr>
<tr>
<td>GA/AA</td>
<td>40</td>
<td>195</td>
</tr>
<tr>
<td>MTHFR(N= 397)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C677T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>22</td>
<td>143</td>
</tr>
<tr>
<td>CT</td>
<td>26</td>
<td>158</td>
</tr>
<tr>
<td>TT</td>
<td>9</td>
<td>39</td>
</tr>
<tr>
<td>CT/TT</td>
<td>35</td>
<td>197</td>
</tr>
<tr>
<td>MTHFR(N= 397)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1298C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>29</td>
<td>168</td>
</tr>
<tr>
<td>AC</td>
<td>23</td>
<td>142</td>
</tr>
<tr>
<td>CC</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>AC/CC</td>
<td>28</td>
<td>172</td>
</tr>
</tbody>
</table>

\textsuperscript{1}Adjusting for clinical covariates stage at diagnosis, chemotherapy use, ER-status did not affect point estimates by more than 10%.

\textsuperscript{2}Only baseline DNA included in the analyses.
There was a statistically significant inverse association between the 
*MTRR* G66A polymorphism and breast cancer-specific mortality: carriers of the 
variant alleles (GA/AA) had reduced risk compared to carriers of the homozygous 
wild-type genotype (GG) (HR: 0.40; 95% CI: 0.2-0.9; P-value: 0.04). The 
association remained statistically significant when the heterozygous GA (HR: 
0.38; 95% CI: 0.2-1.0; P-value: 0.05) and homozygous AA (HR: 0.42; 95% CI: 
0.1-1.4; P-value: 0.15) were compared to the GG genotype. Similar significant 
inverse associations were also observed between carriers of the variant alleles of 
the *MTRR* G66A polymorphism and all-cause mortality compared to carriers of 
the homozygous dominant genotype.

A non-significant, positive association was observed between the *MTHFR* 
C677T polymorphism and breast cancer-specific mortality: carriers of the variant 
allele (CT/TT) had increased risk compared to the women who were 
homozygous for the major allele (CC) (HR: 2.17; 95% CI: 0.8-5.9; p-value: 0.13). 
This positive trend for association also was evident when the heterozygous CT 
(HR: 2.22; 95% CI: 0.8-6.3; p-value: 0.13) and homozygous TT (HR: 2.14; 95% 
CI: 0.5-8.9; p-value: 0.29) genotypes were separately compared to the CC 
referent group. Although there was no association between carriers of CT 
genotype and all-cause mortality, there was a positive association among 
carriers of homozygous TT genotype and all-causes mortality (HR: 1.56; 95% CI: 
0.7-3.4; p-value: 0.27).

A non-significant, inverse association with breast cancer mortality also 
was observed among carriers of the variant allele (G) of the *MTR A2756G*
polymorphism. No consistent association was observed for all-cause mortality and \textit{MTR} A2756G genotype. The three other SNPs, \textit{BHMT} A742G, \textit{SHMT1} G1420A, and \textit{FOLR1} G606A, were not statistically, significantly associated with either breast cancer-specific or all-cause mortality in this sample (table 14).

\textbf{Test for confounding:}

To explore if demographic and clinical covariates such as race/ethnicity, BMI, menopausal status, stage at diagnosis, chemotherapy, ER-status confounded the associations between the genes in the folate metabolizing pathway and breast cancer-specific mortality, multivariate Cox proportional hazard analysis models were tested. Models with covariates including dietary intakes of folate and B complex vitamins, smoking status and alcohol consumption were examined. Age-adjusted HRs reported in tables 14 did not change with the incorporation of these covariates of interest in the models. None of these covariates in multivariate Cox proportional hazard analysis models produced a change greater than 10% in the main effects of the SNPs.

\textbf{AIM 2: GENE - GENE INTERACTION}

Gene-gene interaction was significant for two interactions (\textit{BHMT} A742G X \textit{FOLR1} G60A and \textit{MTRR} G66A X \textit{MTHFR} C677T, (see table 15)).
Table 15. Age-adjusted breast cancer-specific mortality stratified by genotypes in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study (n= 446).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Deaths N</th>
<th>Censored N</th>
<th>Hazard Ratio (^1) 95% CI</th>
<th>P-value</th>
<th>Deaths N</th>
<th>Censored N</th>
<th>Hazard Ratio (^1) 95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BHMT A742G</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>FOLR1 G606A (GG)</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>6</td>
<td>95 94</td>
<td>1.00</td>
<td>0.03</td>
<td>2</td>
<td>101 98</td>
<td>0.33 0.1-1.6</td>
<td>0.17</td>
</tr>
<tr>
<td>AG/GG</td>
<td>3</td>
<td>107 97</td>
<td>0.44 0.1-1.8</td>
<td>0.25</td>
<td>11</td>
<td>121 92</td>
<td>1.42 0.5-3.8</td>
<td>0.49</td>
</tr>
<tr>
<td>P for interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td><strong>MTHFR C677T</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>MTHFR C677T (CC)</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>1</td>
<td>53 98</td>
<td>1.00</td>
<td>0.87</td>
<td>8</td>
<td>47 84</td>
<td>3.96 1.2-13.2</td>
<td>0.02</td>
</tr>
<tr>
<td>GA/AA</td>
<td>4</td>
<td>109 95</td>
<td>0.89 0.2-3.6</td>
<td>0.87</td>
<td>6</td>
<td>171 96</td>
<td>0.87 0.2-3.1</td>
<td>0.82</td>
</tr>
<tr>
<td>P for interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.03</td>
</tr>
</tbody>
</table>

\(^1\) Adjusting for clinical covariates stage at diagnosis, chemotherapy use, ER-status did not affect point estimates by more than 10%.

\(^*\) Only baseline DNA included in the analyses.
For the model examining interaction between BHMT A742G and FOLR1 G606A, women with AA genotype for BHMT A742G and the GG genotype for FOLR1 G606A were considered the referent group. Women with at least one variant allele for BHMT A742G (GA/GG) and FOLR1 G606A (GG) (HR: 0.44; 95% CI: 0.1-1.8; p-value: 0.25) and those with the AA genotype for BHMT A742G and GA/AA for FOLR1 G606A (HR: 0.33; 95% CI: 0.1-1.6; p-value: 0.17) had an inverse association with breast cancer-specific mortality (Figures 13 and 14). Women with at least one variant allele for both FOLR1 G606A polymorphism (GA/AA) and BHMT A742G (AG/GG) had a positive association with breast cancer-specific mortality rates (HR: 1.42; 95% CI: 0.5-3.8; p-value: 0.49). The P value for interaction was 0.03. There was synergistic effect of the variant alleles of the two polymorphisms on breast cancer-specific mortality.
For the model investigating the interaction between $MTRR$ G66A and $MTHFR$ C677T, women carrying the genotypes GG for $MTRR$ G66A and CT/TT for $MTHFR$ C677T were almost four times more likely to die due to breast cancer.
(HR: 3.96; 95% CI: 1.2-13.2; p-value: 0.02) compared to women who were homozygous carriers of both major alleles (GG for MTRR G66A and CC for MTHFR C677T) (figure 15). There was non-significant inverse association with breast cancer-specific mortality for women carrying the genotypes GA/AA for MTRR G66A and CC for MTHFR C677T (figure 16). Non-significant associations were also observed for those with at least one variant genotype for both polymorphisms (GA/AA for MTRR G66A and CT/TT for MTHFR C677T) compared to women with GG for MTRR G66A and CC for MTHFR C677T. In this multiplicative model P for interaction was 0.08. There was borderline antagonistic effect of the variant allele on breast cancer-specific mortality.
Figure 15: Kaplan-Meier survival curves among carriers of the *MTRR* G66A, GG genotype stratified by *MTHFR* C677T polymorphism, HEAL study, New Mexico (N= 446)

Figure 16: Kaplan-Meier survival curves among carriers of the *MTRR* G66A, GA/AA genotype stratified by *MTHFR* C677T polymorphism, HEAL study, New Mexico (N= 446)
AIM 3: GENE - ENVIRONMENT INTERACTION

Interaction with chemotherapy:

To explore if chemotherapy influences the effect of the genes in the folate metabolizing pathway on breast cancer-specific mortality, the associations between the polymorphisms and breast cancer mortality stratified by chemotherapy was examined in table 16.
Table 16: Age-adjusted breast cancer-specific mortality stratified by chemotherapy and genotype in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study (n=446).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Deaths N</th>
<th>Censored N</th>
<th>Censored %</th>
<th>Hazard Ratio(^\dagger) 95% CI</th>
<th>P-value</th>
<th>Deaths N</th>
<th>Censored N</th>
<th>Censored %</th>
<th>Hazard Ratio(^\dagger) 95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHMT A742G (N=446)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MTR A2756G (N=443)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>4</td>
<td>146</td>
<td>97</td>
<td>1.00</td>
<td>0.20</td>
<td>4</td>
<td>50</td>
<td>93</td>
<td>3.26 0.8-13.8</td>
<td>0.11</td>
</tr>
<tr>
<td>AG/GG</td>
<td>4</td>
<td>180</td>
<td>98</td>
<td>0.81 0.2-3.3</td>
<td>0.77</td>
<td>10</td>
<td>48</td>
<td>83</td>
<td>8.59 2.4-30.7</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>BHMT A742G (N=446)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MTR G66A (N=441)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>6</td>
<td>208</td>
<td>97</td>
<td>1.00</td>
<td>0.64</td>
<td>8</td>
<td>60</td>
<td>88</td>
<td>4.06 1.3-12.4</td>
<td>0.01</td>
</tr>
<tr>
<td>AG/GG</td>
<td>2</td>
<td>118</td>
<td>98</td>
<td>0.52 0.1-2.5</td>
<td>0.41</td>
<td>5</td>
<td>36</td>
<td>88</td>
<td>3.99 1.2-13.6</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>MTRR G1420A (N=443)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHMT A1098G (N=443)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>4</td>
<td>87</td>
<td>95</td>
<td>1.00</td>
<td>0.99</td>
<td>6</td>
<td>28</td>
<td>82</td>
<td>3.70 1.0-13.2</td>
<td>0.04</td>
</tr>
<tr>
<td>GA/AA</td>
<td>4</td>
<td>240</td>
<td>98</td>
<td>0.29 0.1-1.1</td>
<td>0.07</td>
<td>7</td>
<td>69</td>
<td>91</td>
<td>1.87 0.5-6.4</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>SHMT A1098G (N=443)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>5</td>
<td>157</td>
<td>97</td>
<td>1.00</td>
<td>0.05</td>
<td>5</td>
<td>47</td>
<td>90</td>
<td>3.73 1.0-13.9</td>
<td>0.05</td>
</tr>
<tr>
<td>GA/AA</td>
<td>3</td>
<td>170</td>
<td>98</td>
<td>0.56 0.1-2.3</td>
<td>0.44</td>
<td>9</td>
<td>50</td>
<td>85</td>
<td>6.49 1.9-21.9</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

\(^\dagger\) Adjusting for clinical covariates stage at diagnosis, chemotherapy use, ER-status did not affect point estimates by more than 10%;

\(^\dagger\) Only baseline DNA included in the analyses.
Table 16 continued: Age- adjusted breast cancer-specific mortality stratified by chemotherapy and genotype in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study (n= 446).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Deaths</th>
<th>Censored</th>
<th>Hazard Ratio</th>
<th>P-value</th>
<th>Deaths</th>
<th>Censored</th>
<th>Hazard Ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td>Chemotherapy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FOLR1 G606A (N= 445)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>4</td>
<td>157 98</td>
<td>1.00</td>
<td></td>
<td>5</td>
<td>45 90</td>
<td>4.63 1.2-18.1</td>
<td>0.02</td>
</tr>
<tr>
<td>GA/AA</td>
<td>4</td>
<td>169 98</td>
<td>0.89 0.2-3.6</td>
<td>0.87</td>
<td>9</td>
<td>53 85</td>
<td>7.07 2.0-24.5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>P for interaction</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTHFR C677T (N= 397)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>1</td>
<td>175 98</td>
<td>1.00</td>
<td></td>
<td>4</td>
<td>32 89</td>
<td>7.23 1.5-33.8</td>
<td>0.01</td>
</tr>
<tr>
<td>CT/TT</td>
<td>6</td>
<td>166 97</td>
<td>2.00 0.5-8.0</td>
<td>0.33</td>
<td>9</td>
<td>51 85</td>
<td>10.56 2.7-41.4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>P for interaction</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTHFR A1298C (N= 397)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>4</td>
<td>192 97</td>
<td>1.00</td>
<td></td>
<td>6</td>
<td>42 88</td>
<td>4.88 1.5-16.0</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AC/CC</td>
<td>3</td>
<td>149 98</td>
<td>0.63 0.2-2.5</td>
<td>0.51</td>
<td>7</td>
<td>41 85</td>
<td>5.52 1.7-17.6</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>P for interaction</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1Adjusting for clinical covariates stage at diagnosis, chemotherapy use, ER-status did not affect point estimates by more than 10%;

2Only baseline DNA included in the analyses.
In these analyses, the referent group was women homozygous for the major allele and did not receive chemotherapy as part of their treatment following breast cancer diagnosis. For BHMT A742G, breast cancer-specific mortality was higher among those with AA genotype who received chemotherapy (HR: 3.26; 95% CI: 0.8-13.8; p-value: 0.11) compared to those with the same genotype who did not receive chemotherapy (table 17). Among women who did not receive chemotherapy, the AG/GG genotype had a non-significant inverse association with breast cancer-specific mortality (HR: 0.81; 95% CI: 0.2-3.3; p-value: 0.77) compared to carrying the AA genotypes (figure 18). Women with the same genotype who received chemotherapy had significantly increased breast cancer-specific mortality rates (HR: 8.59; 95% CI: 2.4-30.7; p-value: <0.01) compared to not carrying the variant allele who did not receive chemotherapy suggesting a possible synergistic effect of the polymorphism and chemotherapy on breast cancer-specific mortality (p for interaction: 0.21).
Figure 17: Kaplan-Meier survival curves among carriers of the BHMT A742G, AA genotype stratified by chemotherapy, HEAL study, New Mexico (N= 446)

Figure 18: Kaplan-Meier survival curves among carriers of the BHMT A742G, AG/GG genotype stratified by chemotherapy, HEAL study, New Mexico (N= 446)
For *SHMT* G1420A polymorphism, women with GG genotype, who received chemotherapy were at increased risk of breast cancer-specific mortality (HR: 3.73; 95% CI: 1.0-13.9; p-value: 0.05) compared to women without chemotherapy (figure 19). Women with the variant alleles of the polymorphisms (GA/AA) who did not receive chemotherapy had inverse association with breast cancer-specific mortality (HR: 0.56; 95% CI: 0.1-2.3; p-value: 0.44) (figure 20) while those with the same genotype who received chemotherapy were at increased breast cancer-specific mortality rates (HR: 6.49; 95% CI: 1.9-21.9; p-value: <0.01) compared to women with the homozygous dominant genotype (GG) without chemotherapy suggesting a borderline synergistic effect of *SHMT* G1420A polymorphism and chemotherapy on breast cancer-specific mortality (p for interaction: 0.21).

No significant interactions with chemotherapy were observed with the *MTR, MTR, and FOLR1* polymorphisms and breast cancer-specific mortality.
Figure 19: Kaplan-Meier survival curves among carriers of the *SHMT* G1420A, GG genotype stratified by chemotherapy, HEAL study, New Mexico (N= 446)

Figure 20: Kaplan-Meier survival curves among carriers of the *SHMT* G1420A, GA/AA genotype stratified by chemotherapy, HEAL study, New Mexico (N= 446)
Interaction with ER-status:

To investigate if ER-status modifies the effect of the genes in the folate metabolizing pathway on breast cancer-specific mortality, the associations between the polymorphisms and breast cancer-specific mortality stratified by ER status were examined in table 17.
Table 17: Age-adjusted breast cancer-specific mortality stratified by ER status and genotype in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study (n= 446).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Deaths</th>
<th>Censored</th>
<th>Hazard Ratio&lt;sup&gt;1&lt;/sup&gt;</th>
<th>P-value</th>
<th>Deaths</th>
<th>Censored</th>
<th>Hazard Ratio&lt;sup&gt;1&lt;/sup&gt;</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N (%)</td>
<td>95% CI</td>
<td></td>
<td>N</td>
<td>N (%)</td>
<td>95% CI</td>
<td></td>
</tr>
<tr>
<td><strong>ER-Positive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EHMT A742G (N= 313)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>3</td>
<td>120 90</td>
<td>1.00</td>
<td></td>
<td>3</td>
<td>26 90</td>
<td>4.33 0.9-21.5</td>
<td>0.07</td>
</tr>
<tr>
<td>AG/GG</td>
<td>7</td>
<td>127 95</td>
<td>2.18 0.6-8.4</td>
<td>0.26</td>
<td>4</td>
<td>23 85</td>
<td>6.21 1.3-28.7</td>
<td>0.01</td>
</tr>
<tr>
<td>P for interaction</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MTR A2756G (N= 313)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>8</td>
<td>154 94</td>
<td>1.00</td>
<td></td>
<td>2</td>
<td>29 94</td>
<td>1.09 0.2-5.1</td>
<td>0.91</td>
</tr>
<tr>
<td>AG/GG</td>
<td>2</td>
<td>93 98</td>
<td>0.37 0.1-1.7</td>
<td>0.19</td>
<td>4</td>
<td>20 83</td>
<td>2.92 0.9-9.7</td>
<td>0.08</td>
</tr>
<tr>
<td>P for interaction</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MTRR G66A (N= 313)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>5</td>
<td>62 91</td>
<td>1.00</td>
<td></td>
<td>2</td>
<td>18 90</td>
<td>1.06 0.2-5.3</td>
<td>0.95</td>
</tr>
<tr>
<td>GA/AA</td>
<td>4</td>
<td>185 98</td>
<td>0.23 0.1-0.8</td>
<td>0.02</td>
<td>5</td>
<td>31 86</td>
<td>1.51 0.5-5.1</td>
<td>0.50</td>
</tr>
<tr>
<td>P for interaction</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SHMT G1420A (N= 313)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>4</td>
<td>125 97</td>
<td>1.00</td>
<td></td>
<td>4</td>
<td>25 86</td>
<td>4.41 1.1-17.9</td>
<td>0.03</td>
</tr>
<tr>
<td>GA/AA</td>
<td>6</td>
<td>122 95</td>
<td>1.51 0.4-5.4</td>
<td>0.52</td>
<td>3</td>
<td>24 89</td>
<td>3.60 0.8-16.6</td>
<td>0.09</td>
</tr>
<tr>
<td>P for interaction</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup>Adjusting for clinical covariates stage at diagnosis, chemotherapy use, ER-status did not affect point estimates by more than 10%.

<sup>2</sup>Only baseline DNA included in the analyses.
Table 17 continued: Age-adjusted breast cancer-specific mortality stratified by ER status and genotype in the New Mexico samples of the Health, Eating, Activity, and Lifestyle Study (n=446).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Deaths</th>
<th>Censored</th>
<th>Hazard Ratio 1</th>
<th>P-value</th>
<th>Deaths</th>
<th>Censored</th>
<th>Hazard Ratio 1</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>95% CI</td>
<td></td>
<td>N</td>
<td>N</td>
<td>95% CI</td>
<td></td>
</tr>
<tr>
<td><strong>ER-Positive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOLR1 G606A (N=313)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>3</td>
<td>119</td>
<td>0.98</td>
<td>1.00</td>
<td>3</td>
<td>25</td>
<td>0.90-21.5</td>
<td>0.07</td>
</tr>
<tr>
<td>GA/AA</td>
<td>7</td>
<td>128</td>
<td>95</td>
<td>2.19</td>
<td>4</td>
<td>24</td>
<td>3.04-7.3</td>
<td>0.01</td>
</tr>
<tr>
<td>P for interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.67</td>
</tr>
<tr>
<td><strong>MTHFR C677T (N=397)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>2</td>
<td>131</td>
<td>97</td>
<td>1.00</td>
<td>2</td>
<td>17</td>
<td>0.70-19.8</td>
<td>0.14</td>
</tr>
<tr>
<td>CT/TT</td>
<td>6</td>
<td>122</td>
<td>95</td>
<td>1.60</td>
<td>5</td>
<td>26</td>
<td>1.5-21.1</td>
<td>0.01</td>
</tr>
<tr>
<td>P for interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.72</td>
</tr>
<tr>
<td><strong>MTHFR A1298C (N=397)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>3</td>
<td>139</td>
<td>97</td>
<td>1.00</td>
<td>4</td>
<td>23</td>
<td>1.2-17.1</td>
<td>0.03</td>
</tr>
<tr>
<td>AC/CC</td>
<td>5</td>
<td>114</td>
<td>96</td>
<td>1.20</td>
<td>3</td>
<td>20</td>
<td>0.8-15.2</td>
<td>0.08</td>
</tr>
<tr>
<td>P for interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
</tr>
</tbody>
</table>

1 Adjusting for clinical covariates stage at diagnosis, chemotherapy use, ER-status did not affect point estimates by more than 10%.

2 Only baseline DNA included in the analyses.
In these survival analysis models, in the referent group was women with the homozygous dominant genotype with ER-positive tumor. For MTR A2756G, for women with AA genotype, there was no association between women with ER-negative tumors and breast cancer-specific mortality compared to women with ER-positive tumor (figure 21). Figure 21 shows violation of the proportional hazard assumption. Among women who carried at least one variant allele of the polymorphism (AG/GG), those with ER-positive tumor had an inverse association with breast cancer-specific mortality (HR: 0.37; 95% CI: 0.1-1.7; p-value: 0.19) (figure 22), while women with ER-negative tumor had significantly increased risk for breast cancer-specific mortality (HR: 2.92; 95% CI: 0.9-9.7; p-value: 0.08) suggesting multiplicative interaction (p for interaction: 0.11).
For **MTRR** G66A polymorphism, there was no association between women with ER-negative tumor and GG genotype and breast cancer-specific mortality compared to women with the same genotype and ER-positive tumor (figure 23). The Kaplan-Meier curve in figure 23 shows the violation of the proportional hazard assumption. In this sample, among women who carried at least one
variant allele of the polymorphism (GA/AA), there was significant inverse association among women with ER-positive tumor and breast cancer-specific mortality (HR: 0.23; 95% CI: 0.06-0.79; p-value: 0.02) compared to women with GG genotype and ER-positive tumor (figure 24). There was non-significant association between women with ER-negative tumor and GA/AA genotype and breast cancer-specific mortality (HR: 1.51; 95% CI: 0.5-5.1; p-value: 0.50) compared to women with GG genotype and ER-positive tumor suggesting multiplicative interaction (p for interaction: 0.11).
Figure 23: Kaplan-Meier survival curves among carriers of the MTR A2756G, AA genotype stratified by ER-status, HEAL study, New Mexico (N= 446)

Figure 24: Kaplan-Meier survival curves among carriers of the MTR A2756G, AG/GG genotype stratified by ER-status, HEAL study, New Mexico (N= 446)
CHAPTER VI
DISCUSSION

INTRODUCTION

This chapter summarizes the findings of the dissertation project and provides potential interpretations that establish integration between biochemical mechanisms of the folate metabolizing pathway reported in the literature and the statistical findings of the current study. Strengths and limitations of this dissertation are evaluated also.

In this population-based breast cancer cohort with 446 participants from New Mexico, 5% of the study participants died due to breast cancer related complications within the 11.46 years of follow-up. Since 1990 there has been a steady decline in the breast cancer-specific mortality rates mostly due to early detection, advancement in treatment options and decreased incidence. Although the literature on the genes in the folate metabolizing pathway and their influence on breast cancer survival are limited, studies have investigated the associations between MTHFR polymorphisms and breast cancer risk. The genes in the folate metabolizing pathway are responsible for enzymes that maintain DNA methylation, synthesis and repair. Since abnormal DNA synthesis and repair may lead to carcinogenesis and irregularity in DNA methylation may modify response to chemotherapy, genes in the folate metabolizing pathway may influence breast cancer survival.
AIM1: MAIN EFFECTS

In this exploratory study, MTRR G66A had a statistically significant inverse association with breast cancer-specific mortality and all-causes mortality. Similar inverse associations were also observed in the Long Island Breast Cancer study (Breast cancer-specific mortality: HR: 0.77; 95% CI: 0.49-1.23; all-causes mortality: HR: 0.75; 95% CI: 0.52-1.08)\textsuperscript{19}. Non-significant associations were also observed between four SNPs, MTR G66A, BHMT A742G, SHMT1 G1420A, and FOLR1 G606A and breast cancer-specific mortality. In the New Mexico sample, unlike the larger HEAL sample, there was positive association between MTHFR C677T and breast cancer-specific mortality.

AIM 2: GENE - GENE INTERACTION

Significant and borderline gene-gene interactions were observed between FOLR1 G606A and BHMT A742G polymorphisms and between MTRR G66A and MTHFR C677T polymorphisms respectively.

Potential interpretations for gene-gene interactions:

Gene-gene interaction between FOLR1 G606A and BHMT A742G (p for interaction: 0.03) in predicting breast cancer-specific mortality was observed in this sample. One of the reasons for interaction may be related to their respective functions in the folate metabolizing pathway. The role of FOLR1 is to bind with folic acid and its derivatives and facilitate transportation of 5-methylTHF. The role of BHMT enzyme is the re-methylation of homocysteine to methionine\textsuperscript{47}, a process which uses 5-methylTHF transported by FOLR1. Variations in FOLR1 may influence the levels of 5-methylTHF, while variation in BHMT polymorphisms
may compromise the conversion of homocysteine to methionine. The plausibility of this argument is supported by the synergistic effect of the two polymorphisms of these genes in predicting breast cancer survival. Hence, each of these polymorphisms may modify the effect of the other in predicting breast cancer-specific mortality.

In this exploratory study, a statistically significant interaction between MTRR G66A and MTHFR C677T (p for interaction: 0.08) in predicting breast cancer-specific mortality was observed also. Biochemical studies suggest that MTHFR acts as a catalyst in the conversion of 5, 10-methylHF to 5-methylTHF, which serves as a donor for the re-methylation of homocysteine to methionine\textsuperscript{9}. Methionine synthase reductase activates methionine synthase by acting as a carbon donor for homocysteine re-methylation and substitutes isoleucine with methionine\textsuperscript{43}. Variations in MTHFR and MTRR polymorphisms may, therefore, influence homocysteine levels, although in different directions. Since studies are inconsistent about the direction of the association between MTRR polymorphism and homocysteine levels\textsuperscript{43} and it is also unknown if homocysteine levels are associated with breast cancer-specific mortality in this population, the antagonistic effect observed in this exploratory study should be interpreted with caution. In spite of low to moderate power of this study (tables 5 and 6, chapter III) women carrying the variant alleles for both polymorphism had statistically significant reduced breast cancer-specific mortality rates compared to women carrying the homozygous dominant genotypes for both polymorphism while significantly increased breast cancer-specific mortality was observed among
women carrying variant allele for *MTHFR* C677T and homozygous GG genotype for *MTRR* G66A.

**AIM 3: GENE - ENVIRONMENT INTERACTION**

Although it did not reach statistical significance, chemotherapy modified the effects of *BHMT* A742G and *SHMT1* G1420A. In this sample, there were also borderline significant interactions between ER-status and *MTRR* G66A polymorphism and *MTR* A2756G polymorphism.

**Potential interpretations for gene-environment interactions:**

In this sample, women carrying the AG/GG genotype of *BHMT* A742G polymorphism who received chemotherapy may have elevated rates of breast cancer related deaths. The apparent synergistic effect of chemotherapy with *BHMT* A742G polymorphism may be supported by the following arguments. Two common chemotherapeutic agents 5 FU and MTX affect DNA synthesis and down-regulate cell replication. 5'-formylTHF, the precursor of 5, 10-methylenetetrahydrofolate potentiates the cytotoxic effects of 5-FU by stabilizing the inhibition of 5, 10-methylenetetrahydrofolate thymidylate synthase- 5FdUMP. MTX inhibits dihydrofolate reductase, decreases intracellular 5,10-methyleneTHF levels for the synthesis of thymidylate and inhibits the synthesis of purine. Although not directly involved in folate metabolism, BHMT helps in the conversion of homocysteine to methionine by donating carbon when the re-methylation process is compromised due to genetic and/or dietary factors. It is also suggested to be associated with homocysteine levels. Although future research should confirm the findings of this exploratory study, it is plausible that since chemotherapy may influence DNA
synthesis and repair and the BHMT A742G polymorphism may affect the methylation process, the two simultaneous processes may result in the elevation of the breast cancer-specific mortality.

In this study, there was also borderline interaction between chemotherapy and SHMT G1420A polymorphism. It has been suggested that carriers of the minor allele of SHMT G1420A polymorphism may decrease the production of the 5-10-methylTHF for pyrimidine synthesis which may reduce the levels of DNA methylation through negative feedback resulting in decreased 5-methylTHF production and accumulation of THF. The common chemotherapeutic agents 5FU and MTX interferes with the levels of 5-10-methylTHF and increases their cytotoxic effects resulting in abnormal DNA synthesis. Owing to the synergistic effect of the variant allele and the effect of chemotherapy women who received chemotherapy and carried at least one variant allele of SHMT G1420A polymorphism may have increased breast cancer-specific mortality rates compared to carriers of the homozygous dominant genotype GG who did not receive chemotherapy.

Studies have suggested that DNA methylation may be associated with inhibition of ER expression. Binding of estrogen to the ER is thought to stimulate mammary cell proliferation, which may lead to decrease in cell division, DNA replication and mutation. Although studies have reported the presence of association between MTR A2756G and plasma homocysteine in humans, the results are inconsistent for the levels of homocysteine across genotypes of A2756G. Like the MTR polymorphism, although studies are consistent about
the association between homocysteine levels and MTRR G66A there is
inconsistency in the literature regarding the direction of the association
Thus, variations in either of these polymorphisms may compromise DNA
methylation which could be associated with ER-induced carcinogenesis. Future
studies in this area may explain the synergistic of these polymorphisms and ER
status.

STRENGTHS

One of the strengths of this study is that although the sample for this
population based case cohort is relatively small it has a maximum follow-up time
of 11.5 years. Secondly, no other study has so far investigated the interactions
among the genes in the folate metabolizing pathway and breast cancer-specific
mortality. Interactions of the genes in the folate metabolizing pathway with
chemotherapy and ER-status were also investigated. Since the enzymes
associated with these genes in the folate metabolizing pathway influence each
other and are known in the biochemical literature to be influenced by prognostic
markers of breast cancer, it is important to investigate the effects of these
interactions in predicting breast cancer-specific mortality. Hence, the study is an
important contribution to the literature on folate metabolizing pathway and breast
cancer-specific mortality.

LIMITATIONS

One of the potential limitations of this dissertation project is its small
sample size. However, the study participants were followed for a maximum of
11.5 years which helped enhance the power of the study to a great extent.
Secondly, in this study only the samples from New Mexico site of the HEAL study were used. The sample consisted of 22% Hispanic and 78% non-Hispanic white ethnicity. There was, however, limited power to investigate ethnic-specific associations. Third, the study also had limited power to investigate the effects of folate and B-complex vitamins on the association of these genes in the folate metabolizing pathway and breast cancer mortality. Although interactions with chemotherapy and genes in the folate metabolizing pathway in predicting breast cancer-specific mortality rates were investigated, effect modifications with individual chemotherapeutic regimens and breast cancer-specific mortality could not be studied due to lack of power. Potential three-way interactions may exist between genes in the folate metabolizing pathway in predicting breast cancer-specific mortality rates, but was unable to investigate due to limited sample size in this study.

Fourth, all analyses in this dissertation were based on dominant models. It would be of interest to know the strengths of the associations between genes in the folate metabolizing pathway and breast cancer-specific mortality and all-causes mortality when other models are used but owing to limited power in this cohort, results of other model types would be unstable.

Fifth, the current study did not investigate three other genes that influence folate metabolism: dihydrofolate reductase (DHFR) and thymidylate synthetase (TYMS or TS) and replication factor C1 (RFC1), which were studied in the Long Island Breast Cancer Study and have been investigated by researchers who have studied the pathway in relation to etiology of other cancers. TYMS and
DHFR serve important roles in folate metabolizing pathway. TYMS catalyzes the conversion of dUMP to dTMP, which is important for DNA synthesis and repair\textsuperscript{115}, while 5,10-methylene-THF, a substrate of TYMS, is obtained from thymidine synthesis, purine synthesis, and DNA methylation\textsuperscript{21}. Commonly used chemotherapeutic agents, 5-FU and MTX, affects TYMS and DHFR and thereby block DNA synthesis and cell replication\textsuperscript{19}. However, unlike the SNPs corresponding to MTHFR, BHMT, MTR, MTRR and SHMT genes, where there was only one nucleotide change among the two variants of the polymorphs, TYMS, with repeats at 5'-untranslated region (5'-UTR) and insertion-deletion at 3' UTR and DHFR, which is a 19 bp deletion, requires more advanced laboratory capabilities than were available for this research and therefore were not included. Like FOLR1, RFC1 is responsible for transportation of 5-methylTHF from plasma to cytosol\textsuperscript{116}. Although a commonly investigated SNP, A80G (rs1051266) polymorph of RFC1, was not investigated in this dissertation because its population allele frequency (Minor allele frequency: 0.45), was beyond the range that was set a priori for a SNP to be eligible for investigation in this study (refer to chapter 3 for eligibility criteria for SNP selection).

Sixth, in the preliminary study the association between MTHFR polymorphisms and breast cancer survival (see Chapter IV for details) was investigated using data from all three HEAL Study sites (N= 772). Unlike the dissertation project where both baseline and follow-up data were used, only baseline DNA was used while genotyping the MTHFR polymorphisms. In the preliminary study sample, 41 out of the 107 women who died during the follow-up
were from the New Mexico site. For the current project, there was insufficient archival DNA left for genotyping 22 out of the 41 women leading to a systematic selection bias in the sample used in this dissertation. The effect of this bias is reflected in the point estimates of the MTHFR polymorphisms in the two studies (compare table 9, chapter IV and table 14, chapter V). In the larger sample, carriers of the variant alleles of MTHFR C677T (CT/TT: HR: 0.76; 95% CI: 0.5-1.10) and MTHFR A1298C (AC/CC: HR: 0.71; 95% CI: 0.48-1.05) polymorphisms had an inverse association with breast cancer-specific mortality. The direction of the association between MTHFR C677T and breast cancer-specific mortality reverses when the analysis is restricted to the New Mexico sample (CT/TT: HR: 2.17; 95% CI: 0.8-5.9).

Lastly, the current research also did not take into consideration competing risks which may have influenced mortality. Previous work has suggested that the genes in folate metabolizing pathway influence the etiology of colon cancer\textsuperscript{21,22,76,93}. It is, therefore, intriguing to test what percentage of the deaths due to cancer other than breast cancer in this sample could be attributable to colon cancer and if they serve as confounding factors in the association between the genes in the folate metabolizing pathway and breast cancer survival. Studies have shown that higher concentration of homocysteine, an amino acid involved in SAM synthesis, elevates the risks of CVD\textsuperscript{94,95} and MTR, BHMT, MTRR, and MTHFR catalyzes the conversion of homocysteine to methionine by remethylation (chapter II, figure 1). Homocysteine is excreted by the kidney in moderate amounts\textsuperscript{96} but polymorphisms of genes involved in
homocysteine metabolisms could increase the plasma concentration of homocysteine which is suggestive of increasing the risk for CVD. Since the genes in the folate metabolizing pathway may be involved in altering CVD mortality, it may be considered an important competing risk against the association between genes in folate metabolizing pathway and breast cancer mortality. In spite, of its importance in the literature on the folate metabolizing pathway, it could not be investigated in the current dissertation project owing to limited power.

CONCLUSION

To summarize, genes in the folate metabolizing pathway may play an important role in breast cancer-specific mortality. Gene-gene interactions and interactions between prognostic markers like chemotherapy and ER-status and the genes in the folate metabolizing pathway may influence breast cancer-specific mortality. Because folate metabolism is a complex interconnected mechanism, future research should further investigate the findings of this study and help integrate the biochemical mechanisms of the folate metabolizing pathway with statistical findings of Genetic Epidemiological studies. It will also be valuable to the literature on the folate metabolizing pathway, if future studies investigate the effects of these genes on disease-free survival following breast cancer diagnosis and the role of competing risks on all-causes mortality. Larger cohort and randomized control trails should investigate the finding of the current study in relation to breast cancer survival.
REFERENCES

folate levels and breast cancer susceptibility: a case-control study in Taiwan. *Carcinogenesis* 2006;11:2295-300.


26. Xu X GM, Wetmur JG, Bradshaw PT, Teitelbaum SL, Neugut AI, Santella RM, Chen J. B-vitamin intake, one-carbon metabolism, and survival in a


based case-control study and meta-analyses. *Int J Cancer* 2007;120:2696 - 703.


62. Chemotherapy Medicines. 7 East Lancaster Avenue, 3rd Floor Ardmore, PA 19003: Breastcancer.org.

63. Chemotherapy Principles: An Indepth Discussion of the Techniques and Its Role in Cancer Treatment: American Cancer Society.


APPENDIX I

LIST OF ABBREVIATIONS

BHMT: betaine-homocysteine-methyltransferase
BMI: Body Mass Index
CVD: Cardiovascular Disease
DHF: dihydrofolate
DHFR: dihydrofolate reductase
DNA: Deoxyribonucleic acid
dTMP: deoxythymidine monophosphate
dUMP: deoxyuridine monophosphate
ER: estrogen receptor
FOLR1: Folate receptor 1
HR: Hazard Ratio
HEAL: Health, Eating, Activity and Lifestyle study
MTHFR: methylenetetrahydrofolate reductase
MTRR: 5-methyltetrahydrofolate-homocysteine methyltransferase reductase
MTR: 5-methyltetrahydrofolate-homocysteine
MTX: Methotrexate
PR: progesterone receptor
RFC: reduced folate carrier
RT-PCR: real-time polymerase chain reaction
SAH: S-adenosylhomocysteine
SAM: S-adenosylmethionine
SHMT1: serine hydroxymethyltransferase 1
SNP: single nucleotide polymorphisms
THF: tetrahydrofolate
TYMS: thymidylate synthetase
5-FU: fluoropyrimidine
APPENDIX II

INSTITUTIONAL REVIEW BOARD- APPROVAL NOTICE

UNIVERSITY OF LOUISVILLE
It's Happening Here.

Human Subjects Protection Program Office
MedCenter One – Suite 202
501 E. Broadway
Louisville, KY 40222-1798
Office: 502.852.3198 Fax: 502.852.2184

August 17, 2011

Richard Baumgartner, Ph.D.
Department of Epidemiology & Population Health
486 East Gray Street, Room 227A
Louisville, KY 40292

RE: IRB# 319.05/ Health, Eating, Activity and Lifestyle (HEAL) Study

Dear Dr. Baumgartner:

The following items for the aforementioned study have been received by the Human Subjects Protection Program Office and approved by the chair of the Institutional Review Board (IRB) through the expedited review procedure according to 45 CFR 46.110(B):

- Amendment #11, dated 05/01/2011

The modifications include:

- The addition of Ratan Das to the study as key personnel.
- Within an established cohort study, specimens from approximately 800 eligible women will be used (includes all 5 sites). This research includes carrying out DNA extraction and re-sequencing, as well as, conduct of assays focused on approximately 8 specific polymorphisms conducted on a PCR machine. The purpose of the assay work is to meet aims included in the doctoral dissertation proposal of Nandita Das.

Training for this individual has been verified.

This action will be reported promptly to the IRB at a scheduled full Board meeting. Please send all inquiries and electronic revised/requested items to our office email address at hsppofc@louisville.edu.

Sincerely,

Julie L. Goldman, MD
Vice Chair, Biomedical Institutional Review Board
JLG@lku
APPENDIX III

TAQMAN SNP GENOTYPING

RT-PCR uses TaqMan SNP genotyping assay which includes two PCR primers specific to the locus under investigation that flanks the SNP, and two oligonucleotide TaqMan probes specific to the allele of interest. Each probe has fluorescent dye at the 5' end and a non-fluorescent quencher with a minor groove binder (MGB) at the 3' end. The two fluorescence dyes used here are FAM and VIC. Conventionally the major allele of a SNP is identified by the FAM while the minor allele is identified by the VIC dye. TaqMan genotyping uses a universal master mix to detect the variants of the six SNPs. In a nutshell, during each PCR cycle, the probe which is hydrolyzed to the allele of interest, emits a fluorescent signal when it cleaves to the 5' exonuclease action of the AmpliTaq Gold DNA Polymerase. The primers amplify the locus of interest in the DNA template and the dye specific to the alleles reports the presence or absence of the allele in the DNA sample. Taq polymerase, which is isolated from thermophilic bacteria, is capable of withstanding very high temperatures, hence used for the purpose. The presence of two probes, each specific to an allele of the SNP under investigation and labeled with two dyes, enables two alleles to be detected in a single tube. The following is an animated version of the technique shown in three chronological steps.
Figure 26: Animated representation of the 5’ nuclease TaqMan assay process. Adopted from TaqMan SNP genotyping assay protocol\textsuperscript{113, 117}

**Legends:**
- VIC dye
- FAM dye
- Quencher
- MGB
- Anti TaqGold DNA polymerase

**Step 1: DNA template with the components of TaqMan assay**
Step 2: Annealing the components of TaqMan assay

Step 3: Polymerization
Below is the protocol developed for genotyping the six SNPs:

Steps for Genotyping:

Instructions to make working and reaction plates for genotyping using RT-PCR

Prepare for bench work:

1. wipe the work bench with 0.01X bleach solution
2. lay out a long sterile sheet
3. Wear gloves

Working plate:

1. Dilute 1X TE buffer to 0.1X by adding de-ionized or double distilled water.
2. Label each working plate. Be as complete as possible <WP_Experiment name_SNP ID_plate #>
3. Four designated wells are considered as controls and are filled with de-ionized water. Make sure they are spread out randomly across the plate. (Consult the template for their exact location). Keep their positions constant for all plates for convenience.
4. Get the DNA box out of the freezer
5. Thaw each vial by rubbing it with hand in circular motion
6. Allocate required amount of DNA into the labeled working plate well and dispose the pipette tip (consult excel sheet for DNA volume and concentration)
7. Using another pipette tip, allocate required amount of TE into each well. 
   To determine the amount of DNA and 0.1X TE required to make working plates is derived from the equation:
\[ C_1V_1 = C_2V_2 \]

\[ => V_1 = \frac{C_2V_2}{C_1} \]

Here, \( C_1 \) = Initial DNA concentration

\( V_1 \) = Total final volume of DNA after addition of 0.1X TE

\( C_2 \) = Final DNA concentration

\( V_2 \) = Final volume of DNA and 0.1X TE mixture in each well of the working plate

For example, for DNA with initial concentration of 47ng/μL (\( C_1 \)), we allocate 75μL of 0.1X TE to 2μL of DNA (\( V_1 \)),

\[
\text{Volume of DNA needed} = \frac{[2 \times 75]}{47} \\
\sim 3
\]

\[
\text{Volume of 0.1X TE needed} = 75 - 3 = 72
\]

In other words, we allocated 3μL of DNA (of 47ng/μL concentration) in each plate and added 72μL of 0.1X TE to make our working stock. Here, enough DNA-TE mixture is made as working stock to allow room for evaporation.

8. Mix the DNA and the TE by using the pipette (counting 1-15 help)

9. Repeat step 2-6, 92 times to fill 92 wells for the working plate

10. Seal the plate with adhesive covers (less expensive non-optical works fine) making sure the edges are smoothed tightly.
11. Centrifuge for 1 minute at 1200 rpm if you are ready to prepare reaction plates. Otherwise, cover with aluminum foil and store in refrigerator for 6-8 months. If stored for long check the amounts in each well for evaporation.

Reaction plate:

1. To prepare the reaction mixture for a particular SNP we need to start with mixing required amounts of assay, de-ionized water and master-mix in a reservoir used for multi-channel pipettor. Consult corresponding charts to ensure the exact amount required for one plate times the total number of plates that will be run in the RT-PCR on a particular lab day (table 18). As the name suggest, in a 96 well plate, there are 96 wells. To make reaction mixture for one well, we need

   2X master mix: 10µL
   40X assay: 0.5 µL
   De-ionized water: 4.5µL
   Total: 10µL

However, additional amount of 10-15% made for each plate to account for evaporation. Therefore, after accounting for the excess, for one plate, we made a reaction mixture of the following:

   2X master mix: 1200µL
   40X assay: 60µL
De-ionized water: 540µL
Total: 1800µL

Below is a quick look at the amounts corresponding to each of the three ingredients required to prepare reaction mixtures

Table 18: Amounts of ingredients required to make reaction mixtures:

<table>
<thead>
<tr>
<th># of plates</th>
<th># of wells</th>
<th>additional Amts*</th>
<th>2X Master Mix</th>
<th>40X assay</th>
<th>Water</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>0.5</td>
<td>4.5</td>
<td>15</td>
</tr>
<tr>
<td>1</td>
<td>96</td>
<td>120</td>
<td>1200</td>
<td>60</td>
<td>540</td>
<td>1800</td>
</tr>
<tr>
<td>2</td>
<td>192</td>
<td>220</td>
<td>2200</td>
<td>110</td>
<td>990</td>
<td>3300</td>
</tr>
<tr>
<td>3</td>
<td>288</td>
<td>330</td>
<td>3300</td>
<td>165</td>
<td>1485</td>
<td>4950</td>
</tr>
<tr>
<td>4</td>
<td>384</td>
<td>450</td>
<td>4500</td>
<td>225</td>
<td>2025</td>
<td>6750</td>
</tr>
<tr>
<td>5</td>
<td>480</td>
<td>540</td>
<td>5400</td>
<td>270</td>
<td>2430</td>
<td>8100</td>
</tr>
<tr>
<td>6</td>
<td>576</td>
<td>660</td>
<td>6600</td>
<td>330</td>
<td>2970</td>
<td>9900</td>
</tr>
</tbody>
</table>

The "total" column can be double checked by multiplying column "additional amts" by 15; for example, 330 X 15 = 4950. Stocks can be stored in the refrigerator for no more than 2-3 days.

2. Label each reaction plate. Be as complete as possible <RP_Experiment name_SNP ID_plate #>

3. Using an 8 tip pipettor, allocate required amount of standardized DNA-TE into all 96 wells (Consult the template for each experiment for the exact location of each ID). Dispose the pipette tips after every row.

4. Using an 8 tip pipettor, pipette the master-mix cocktail into the wells, one row at a time. Dispose the tips after every row.

5. Once all the 96 wells are filled, seal the plate with optical adhesive cover

6. Centrifuge the plate for 1 min at 1200 rpm.
7. Open the relevant template in the computer/RT-PCR corresponding to the experiment.

8. In the “Advance Set-up” tab type in the experiment name and number, and select the following tabs:
   - Step-one Plus instrument (96 well plates)
   - Genotype
   - TaqMan reagent
   - Standard (speed)

   The rest of the conditions are pre-set.

9. In the “Plate Set-up” tab, double check the “View plate layout” and “View well table” to make sure the right templates are open for the experiment.

10. Stick the plate into the RT-PCR and let the reaction begin. The reaction last for about 1hr 30 mins to complete.

11. Once the reaction is over check if the controls are correctly detected.

12. Save the output in the computer and the USB drive.

**Thermal cycler:**

13. While one plate is running in the RT-PCR machine, a second one can be inserted in the thermal cycler. Select “SNP assay PCR” and click “run”.

   The reaction last for about 1hr 30 minutes.

14. Once it is over, if the RT-PCR is running a plate or you are running out of time for the day at the lab, it is okay for the plate to be stored in the refrigerator for a day. Just remember to (a) cover it with aluminum foil prior
to storage and (b) centrifuge the plate for 1 min at 1200 rpm prior to running it on the RT-PCR machine.

15. Make sure you have opened the right template for the reaction.

16. In the “Run Methods” tab, open the Graphical view tab, select and delete the following sections:
   - Pre-PCR Reading (Holding stage)
   - Holding Stage
   - Cycling Stage

   Only the final step, called the “Post-PCR Read (Holding Stage)” where the end point of the reaction is processed in the RT-PCR machine.

17. Make sure options “Pre-PCR Reading” and “Amplification” are checked.

18. Insert the plate in the RT-PCR machine and click “run”.

19. Reaction will last for approximately 2 minutes.

20. Once the reaction is over check if the controls are correctly detected.

21. Save the output in the computer and the USB drive.

If there is not enough time on the day reaction plates are prepared, the reaction plates can be stored in the refrigerator (for about 2-3 days) after step 6 is completed. Start by centrifuging (step 6 onwards) the following lab day.
CURRICULUM VITAE

Nandita Das

Department of Epidemiology and Population Health, School of Public Health and Information Sciences, University of Louisville, 485 E. Gray St., Louisville, KY 40202.
Phone: 502-500-6728 (cell) /502-290-1009 (home)
Email: n0das001@louisville.edu or mrsnanditadas@yahoo.com

Education

Ph.D. (2012): Epidemiology and Population Health, University of Louisville
Mentor: Dr. Richard N. Baumgartner, PhD
Area of Specialization: Cancer and Nutrition; Current GPA: 3.9 out of 4.0

Mentor: Dr. Susan E. Kelly, PhD
Area of Specialization: Medicine and Health; GPA: 4.0 out of 4.0

M.A. (2001): English, Department of English and Foreign Languages, Tezpur University, Naapam, Tezpur, Assam, India
Area of Specialization: English Literature, British Drama; Award: 1st Class 2nd Position

B.A. (1999): English Major, Darrang College, Tezpur, Assam, India

Employment History

01/2012-05/2012: Dissertation Completion Award

01/2010- 12/2011: Graduate Assistant, Susan G. Komen trainee
Long-Term Quality of Life study, a follow-up of the New Mexico Women’s Health Study
❖ Data cleaning: LTQOL study
❖ Data analysis: LTQOL study, NMWHS
❖ Report writing and presentation
❖ Subject recruitment using NMTR info
❖ Data collection (telephone interview): LTQOL study
Data management (subject tracking using Oracle software and Microsoft Access, data filing system, data entry, and quality control)
Genotyping using PCR technology: HEAL sample

08/2008-12/2009: Graduate Assistant
Department of Epidemiology and Population Health, University of Louisville
Teaching assistant: Tutoring (Fall 2009), Epidemiological Methods (Fall 2008, Spring 2009)
Research assistant: for Cancer research projects:
  - Data analysis: HEAL study and KCR data
  - Data collection (telephone interview): LTQOL study
  - Data management: LTQOL study

06/2008-07/2008: Program Assistant
Cancer Prevention and Control Program, James Graham Brown Cancer Center, University of Louisville
  - Data analyze: SEER Cancer Registries
  - Write proposals for research on subsequent cancers following first primary cancer of the breast, cervix uteri, prostate, lung, or colorectal

01/2008-05/2008: Research Assistant, (Part-time)
Diabetes Self-Management program, Department of Psychology, University of Louisville
  - Subject recruitment at UofL Hospital
  - Data collection (Mail-in-interview)
  - Data entry

09/2006- 12/2006: Research Coordinator, (Volunteer) Urogynecology Specialists of Kentuckiana, Louisville
  - Subject recruitment at Norton Hospital
  - Data collection
  - Preliminary data analyses

Darrang College, Tezpur, Assam, India
  - Taught English language and phonetics
  - Taught critical thinking and writing skills

Skills

- Excellent leadership and inter-personal skills
- Excellent organizational capacity
- Expertise in multi-tasking and trouble shooting
- Experience in working individually and with multi-disciplinary teams
- Expertise in research and analysis of robust data
Expertise in comprehensive documentation of methodologies and findings of analysis and writing reports
Expertise in presentation of analysis findings

Memberships

Since 2008      American Association for Cancer Research
                Associate member
Since 2008      Society for Epidemiological Research
Since 2006      North Central Sociological Association
2005-2006       Graduate Student Council
                University of Louisville, Louisville, KY
Since 2005      Alpha Kappa Delta
                International Sociological Honor Society

Honors and Awards

Spring, 2012    Dissertation Completion Award
2010           Society of Epidemiological Research (SER) Student Caucus Travel Award to present poster at 2010 SER Annual Meeting, Seattle, WA.
2009           Graduate Teaching Academy, University of Louisville
2008           1st Prize at Research! Louisville for the poster: Home-based Tele-monitoring with or without Disease Management: Impact on Quality of Life and Self Efficacy in Patients with Heart Failure.
2006           Future Faculty Certificate awarded from North Central Sociological Association
2005           Alpha Kappa Delta, International Sociological Honor Society, for the study: Psychosocial Factors of Pregnancy
2001           1st class second position in M.A. (English)
1994- 1996     Special Merit Scholarship, State Level Advisory Committee, Students and Youth Welfare, Government of Assam, India
1994-1995  Junior Combined Merit Scholarship, Government of Assam, India

Community Service

2006- 2010  Adult Tutor, Adult Literacy Program of The Life-Long Learning Center, Down Syndrome of Louisville, Inc.

2006  Note Taker, Disability Resource Center, University of Louisville

Studies in process for presentation and publication

Submitted Presentation:


Paper/Poster Published:


Das, N., Baumgartner, R.N., Pinkston, C.M., Baumgartner, K.B. (Sep 18-21, 2011) Antioxidant intake and breast cancer survival among Hispanic and non-Hispanic white women: the New Mexico Women's Health Study. Poster presentation: 4th AACR conference on Science of Cancer Health Disparities in Race/Ethnicity and the medically underserved and Susan G. Komen Annual PI and Trainee meeting.


Poster presented: Research!Louisville, Louisville, Kentucky. 1st place award for poster presentation.


Manuscript in the process of publication:


Das, N., Baumgartner, R.N., Pinkston, C.M., Baumgartner, K.B. Vitamins and minerals intake and breast cancer survival among Hispanic and non-Hispanic white women: the New Mexico Women’s Health Study.

Das, N., Baumgartner, R.N., Pinkston, C.M., Baumgartner, K.B. Treatment related risk factors of arm lymphedema among Hispanic and non-Hispanic white breast cancer survivors

Das, N., Baumgartner, R.N., Pinkston, C.M., Baumgartner, K.B. Association between antioxidant intake and risk for arm lymphedema among Hispanic and non-Hispanic white breast cancer survivors
Past Research Support:

- Role: **Graduate Research Assistant**
  Susan G. Komen Foundation. Grant No: KG090926
  Long Term Quality of Life, follow-up of New Mexico Women’s Health Study
  University of Louisville, Louisville, Kentucky

- Role: **Graduate Research Assistant**
  National Institute of Health Grant, “Ethnicity, Breast Cancer Recurrence and Long-Term QOL” Grant No: 5R01CA105266

- Role: **Graduate Assistant**
  Department of Epidemiology and Population Health, School of Public Health and Information Sciences

- Role: **Program Assistant**

- Role: **Research Assistant**

Current research projects:

Baumgartner RN and Baumgartner KB (PI): Folate metabolizing pathway and breast cancer survival (Healthy Eating Activity and Lifestyle (HEAL)). Data collection funded by National Cancer Institute, National Institutes of Health, SEER Grant No: N01-PC-05031-20. Additional funding assistance through the Susan G. Komen Foundation, Grant No: KG090926.

Baumgartner KB (PI): Long Term Quality of Life, New Mexico Women’s Health Study. Grant No: 5R01CA105266.