Glycemic Index and Breast Cancer Risk and Phenotype

Master of Science, 2010
Carolyn Greenberg
Graduate Department of Nutritional Sciences
University of Toronto

ABSTRACT

Ecological studies and results from our low-fat, high-carbohydrate dietary intervention trial suggest that different carbohydrates are associated with breast cancer risk in different ways. We examined the association of diet glycemic index (GI), a ranking of carbohydrate containing foods based on their blood glucose raising potential, with breast cancer risk and phenotype. GI was calculated from multiple food records from subjects in our intervention trial using a nested case-control design (220 cases, 440 controls). GI was not associated with risk of total or estrogen receptor positive breast cancer, tumor size or nodal status. GI was strongly positively associated with hormone negative breast cancer. This finding is potentially important as little is known about the etiology of hormone negative breast cancer, which has a worse prognosis than hormone positive breast cancer. However, this finding is based on a small number of cases and should be replicated in a larger sample.
ACKNOWLEDGEMENTS

I would like to acknowledge and thank the people that supported me throughout my journey to complete this project.

I am forever indebted to Dr. Norman Boyd and Dr. Lisa Martin for suggesting that I pursue a master’s degree and for providing me with the opportunity to conduct this study. I gained immensely from the experience and look forward to beginning to apply the knowledge and skills I have acquired.

Dr. Lisa Martin, my supervisor extraordinaire, advised me throughout the course work, execution of the study and on every aspect of the thesis. She is a wonderful teacher and patiently explained scientific and statistical concepts until I fully understood their impact on this project.

Dr. Norman Boyd, principal investigator of the Diet and Breast Cancer Prevention Study and member of my thesis committee, has always focused my attention on the key issues. As an advisor on this project, Dr. Boyd generously gave his time to review my work and provided valuable insights which improved the clarity and organization of the thesis, presentation and defense.

Dr. Thomas Wolever was a member of my thesis committee and primary resource on all matters pertaining to the glycemic index. He greatly enhanced my understanding of the science related to the glycemic index and I deeply appreciate his willingness to answer all my questions, big or small, and his enthusiastic support of my work.

Dr. Michael Archer was the external examiner for my thesis examination. Dr. Archer’s expertise, fairness and calm demeanour made the oral exam a positive and thought provoking experience.

Lorraine Gougeon helped assign the glycemic index values to our nutrient database. Lorraine and I worked side by side for almost one year, researching and debating the assignment of glycemic index values to over 2,700 food codes. Her expertise on food and nutrition, dedication to excellence and sense of humour made working together a pleasure and undoubtedly were a major factor in the success of this project.

Valentina Kriukov, data manager for the Diet and Breast Cancer Prevention Study, diligently prepared and implemented an automated system for calculating the glycemic index
for the food records in this study. I am awed by, and grateful for her unique expertise and infallible work.

Kyong Ok-Yi, a very kind, generous and talented individual who was always there to help. I am very appreciative of her assistance with the formatting of the final draft of my thesis.

I would like to express my gratitude to the Canadian Breast Cancer Foundation – Ontario Region for my fellowship and financial support of this study, and to the Canadian Breast Cancer Research Alliance and the Ontario Ministry of Health for financial support of the Diet and Breast Cancer Prevention Study.

My beautiful daughter, Kayla and my beloved mother, Katherine Ritter, are my inspiration. These very special women taught me that optimism, perseverance and determination can help you reach beyond your comfort zone to conquer almost any challenge.

Finally, I am profoundly grateful to my husband, Joel, my biggest supporter. When I needed to talk, he knew to just listen. When I needed to be cajoled, he gently pushed and put a perfect cup of coffee in my hand. When I needed a boost of confidence, he encouraged me. And when I desperately needed some quiet time, he simply knew and put on the right music to soothe my spirit. For me, there is no better soul mate.
# TABLE OF CONTENTS

ABSTRACT .......................................................................................................................... ii
ACKNOWLEDGEMENTS .................................................................................................. iii
TABLE OF CONTENTS ...................................................................................................... v
LIST OF TABLES ................................................................................................................ viii
LIST OF FIGURES ............................................................................................................ ix
LIST OF ABBREVIATIONS ............................................................................................... x

## CHAPTER 1: Introduction

1.1 Background ...................................................................................................................... 1
1.2 Outline of Thesis ........................................................................................................... 3

## CHAPTER 2: Literature Review

2.1 Overview of Breast Cancer Development ....................................................................... 4
  2.1.1 Normal Breast Tissue .............................................................................................. 4
  2.1.2 Development of Breast Cancer ............................................................................. 5
  2.1.3 Phenotype ............................................................................................................... 6

2.2 Risk Factors for Breast Cancer ...................................................................................... 7
  2.2.1 Overview of Risk Factors for Breast Cancer ........................................................... 7
  2.2.2 Age .......................................................................................................................... 7
  2.2.3 Mammographic Density ......................................................................................... 8
  2.2.4 Family History and Genetic Factors ......................................................................... 9
  2.2.5 Reproductive Factors ............................................................................................. 10
    2.2.5.1 Menstrual Factors .......................................................................................... 10
    2.2.5.2 Pregnancy Factors ......................................................................................... 11
    2.2.5.3 Sex Hormones ............................................................................................... 11
  2.2.6 Height ....................................................................................................................... 12
  2.2.7 Body Weight ........................................................................................................... 13
  2.2.8 Risk Factors by Estrogen Receptor Status ............................................................... 14
  2.2.9 Dietary Factors ....................................................................................................... 14
    2.2.9.1 Alcohol ............................................................................................................ 15
    2.2.9.2 Total Fat ......................................................................................................... 16
    2.2.9.3 Carbohydrate .................................................................................................. 17
    2.2.9.4 Limitations of Observational Studies ............................................................... 19
    2.2.9.5 Clinical Trials of Low Fat, High Carbohydrate Diets ....................................... 20
  2.2.10 Glycemic Index ...................................................................................................... 23
    2.2.10.1 Nutrients, Foods and Diet GI .......................................................................... 24
    2.2.10.2 Glycemic Index, Glycemic Load and Breast Cancer Risk ............................... 24
    2.2.10.3 Limitations related to measurements of Diet GI ............................................. 29
CHAPTER 3: Methods

3.1 The Diet and Breast Cancer Prevention Trial ..........................................................33
3.2 Glycemic Index and Breast Cancer Risk and Phenotype Study Design ..............37
3.3 Data Collection ........................................................................................................37
3.4 Development of the GI Database ...........................................................................38
  3.4.1 Overview of GI Database .................................................................................38
  3.4.2 Source of GI Values .........................................................................................39
  3.4.3 Description of GI Database ............................................................................41
  3.4.4 Assignment of GI Values .................................................................................42
  3.4.5 Considerations in the Assignment of GI Values ..............................................44
3.5 Quality Control .......................................................................................................49
3.6 Calculating Diet GI ..................................................................................................49
3.7 Pilot Study to Compare Automated vs Manual System ......................................50
3.8 Statistical Methods .................................................................................................51
3.9 Candidates Contribution to the Project ..................................................................52

CHAPTER 4: Results

4.1 Descriptive Statistics ............................................................................................53
  4.1.1 Characteristics of Subjects ..............................................................................53
  4.1.2 Nutrient Intake ................................................................................................56
  4.1.2.1 Energy and Macronutrients ........................................................................56
  4.1.2.2 Types of Carbohydrate ..............................................................................56
4.2 Diet GI ....................................................................................................................59
  4.2.1 Comparison of Diet GI by Case Control Status ..............................................61
  4.2.2 Associations of Diet GI by Subject Characteristics .......................................63
  4.2.3 Correlations between Diet GI and Nutrients ................................................67
4.3 Association of Diet GI with Breast Cancer Risk – Conditional Logistic Regression...........................................................................................................71
  4.3.1 Univariate and Multivariate Models with Diet GI and Risk Factors ...............71
  4.3.2 Multivariate Analysis with Diet GI, Risk Factors and Type of CHO ...............73
4.4 Diet GI and Breast Cancer Phenotype in Cases-only Analysis .......................75
  4.4.1 Risk of Hormone Receptor Negative vs Hormone Receptor Positive Breast Cancer .............................................................75
    4.4.1.1. Characteristics of Subjects by Hormone Receptor Status.....75
    4.4.1.2 Association of Diet GI with Breast Cancer Risk by Hormone Receptor Status .............................................................77
  4.4.2 Association of Diet GI with Tumor Size ...........................................81
  4.4.3 Association of Diet GI with risk of Nodal Status ..............................81

4.5 Summary of Main Results .............................................................................85

CHAPTER 5: Discussion

5.1 Overview of Results ..........................................................................................86
5.2 Potential Biological Mechanisms .......................................................................89
5.3 Strengths and Limitations ...............................................................................92
  5.3.1 Study Design .........................................................................................92
  5.3.2 Dietary Data and GI Values .................................................................93
5.4 Conclusions ........................................................................................................97
5.5 Future Work .......................................................................................................98

CHAPTER 6: References ..........................................................................................100
## LIST OF TABLES

Table 2.1 Cohort Studies of Diet GI and Breast Cancer Risk ..............................................27  
Table 2.2 Case-Control Studies of Diet GI and Breast Cancer Risk ...................................28  

Table 4.1 Selected baseline characteristics of subjects in the nested case-control study ....54  
Table 4.2 Energy, weight and macronutrient intake ............................................................57  
Table 4.3 Intake of types of carbohydrate ............................................................................58  
Table 4.4 Associations of diet GI with characteristics of subjects ........................................65  
Table 4.5 Pearson correlations of diet GI with energy and nutrients ....................................69  
Table 4.6 Regression coefficients of diet GI with types of carbohydrate ...............................70  
Table 4.7 Diet GI and breast cancer risk – conditional logistic regression ...............................72  
Table 4.8 Diet GI and breast cancer risk by estrogen receptor status ....................................74  
Table 4.9 Selected baseline characteristics of breast cancer cases .......................................76  
Table 4.10 Diet GI with risk of hormone receptor negative versus positive breast cancer ...79  
Table 4.11 Association of diet GI with log tumor size (log cm) ..............................................83  
Table 4.12 Diet GI with risk of nodal involvement ...............................................................84
LIST OF FIGURES

Figure 1.1 Trends in incidence of breast cancer in selected countries .........................1

Figure 2.1 Age-specific incidence of breast cancer in Canada and Shanghai (2002) ........8
Figure 2.2 Six categories of mammographic density ....................................................9
Figure 2.3 Correlation of percent energy from fat with breast cancer incidence ..........16
Figure 2.4 Correlations of sugar and starch consumption with breast cancer mortality ....18
Figure 2.5 Associations of post-randomization weight and selected nutrients with breast cancer risk: interquartile OR (95% CI) .................................................................22
Figure 2.6 Meta-analysis of glycemic index and breast cancer risk ..............................26
Figure 2.7 Biologic plausibility ..................................................................................31

Figure 3.1 Design of the Diet and Breast Cancer Prevention Trial ...............................33
Figure 3.2 Design for the GI and Breast Cancer Risk and Phenotype Study ..............37
Figure 3.3 Flowchart of steps for calculating diet GI ...................................................40
Figure 3.4 Distribution of NDS codes among type of matches with tested values .....42
Figure 3.5 Correlation of manual and automated method for determining diet GI values ...50

Figure 4.1 Histograms of diet GI among cases and controls .......................................60
Figure 4.2 Box plots of diet GI at (A) baseline and (B) post-randomization and (C) mean of all food records by case-control status .........................................................62
Figure 4.3 Histograms of tumor size data .................................................................82
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACHO</td>
<td>available carbohydrate</td>
</tr>
<tr>
<td>AJCC</td>
<td>American Joint Committee on Cancer</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>CHO</td>
<td>carbohydrate</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>DBCPT</td>
<td>Diet and Breast Cancer Prevention Trial</td>
</tr>
<tr>
<td>EPIC</td>
<td>European Prospective Investigation into Cancer and Nutrition</td>
</tr>
<tr>
<td>ER</td>
<td>estrogen receptor</td>
</tr>
<tr>
<td>FAO/WHO</td>
<td>Food and Agriculture Organization/World Health Organization</td>
</tr>
<tr>
<td>FFQ</td>
<td>food frequency questionnaire</td>
</tr>
<tr>
<td>FSH</td>
<td>follicle-stimulating hormone</td>
</tr>
<tr>
<td>GH</td>
<td>growth hormone</td>
</tr>
<tr>
<td>GI</td>
<td>glycemic index</td>
</tr>
<tr>
<td>HDL-C</td>
<td>high density lipoprotein cholesterol</td>
</tr>
<tr>
<td>HFCS</td>
<td>high fructose corn syrup</td>
</tr>
<tr>
<td>HRT</td>
<td>hormone replacement therapy</td>
</tr>
<tr>
<td>IARC</td>
<td>International Agency for Research on Cancer</td>
</tr>
<tr>
<td>IGF</td>
<td>insulin-like growth hormone</td>
</tr>
<tr>
<td>IGFBP</td>
<td>insulin-like growth hormone binding protein</td>
</tr>
<tr>
<td>IQR</td>
<td>interquartile range</td>
</tr>
<tr>
<td>IQQR</td>
<td>odds ratio across the inter-quartile range</td>
</tr>
<tr>
<td>LH</td>
<td>luteinizing hormone</td>
</tr>
<tr>
<td>MDA</td>
<td>malonaldehyde</td>
</tr>
<tr>
<td>NDS</td>
<td>Nutrient Data System</td>
</tr>
<tr>
<td>OC</td>
<td>oral contraceptives</td>
</tr>
<tr>
<td>OR</td>
<td>odds ratio</td>
</tr>
<tr>
<td>PR</td>
<td>progesterone receptor</td>
</tr>
<tr>
<td>RS</td>
<td>resistant starch</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SE</td>
<td>standard error</td>
</tr>
<tr>
<td>SHBG</td>
<td>sex-hormone binding globulin</td>
</tr>
<tr>
<td>SNPs</td>
<td>single nucleotide polymorphisms</td>
</tr>
<tr>
<td>WHI</td>
<td>Women’s Health Initiative</td>
</tr>
<tr>
<td>WHR</td>
<td>waist to hip ratio</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Background

Invasive breast cancer is the most common malignancy in women in the world. Worldwide more than 1.3 million new cases of female breast cancer are diagnosed each year accounting for over 1/5th of the cancers diagnosed in females. According to a recent report from the International Agency for Research on Cancer (IARC), it is now the most common cancer in both developed and developing regions with around 690,000 new cases annually in each region (1). In Canada we expect 22,700 new cases of breast cancer to be diagnosed in 2009 and breast cancer is the second leading cause of cancer mortality among women after lung cancer (2).

![Figure 1.1 Trends in incidence of breast cancer in selected countries](image)

Source: Globocan 2008 (IARC)

Incidence rates of breast cancer vary greatly between countries. The highest incidence rates occur in North America, Northern and Western Europe, Australia, New Zealand, Uruguay and Argentina (3) where the standardized rates are about three times higher than in
lower risk countries in Asia (see figure 1.1) (3). Although incidence rates are increasing over time, women in low risk countries still have a much lower incidence rate of breast cancer compared to women in high risk countries. Migrants from Asia to the West acquire the breast cancer incidence of their adopted country within one or two generations (4-6).

International variations in breast cancer risk, migrant studies, ecological studies and animal studies suggest that environmental factors, such as diet may be important in the development of breast cancer. Although the nature of these factors is uncertain, international differences in breast cancer incidence are strongly positively correlated with per capita estimates of fat, protein, animal protein and sugar intake, and negatively correlated with intake of cereals (7). Animal experimental data show that higher fat intake promotes the development of mammary tumors in rats and mice. The effect of dietary fat on mammary tumor incidence is independent of caloric intake and the magnitude of the effect is estimated to be about 2/3 that of caloric intake (8). These and other data led us to carry out a long-term randomized controlled trial in 4,690 women at increased risk for breast cancer because of extensive mammographic density, to determine if the incidence of breast cancer could be reduced by a low fat, high carbohydrate (CHO) diet.

We did not find a significant difference in breast cancer incidence between the control and low fat, high CHO group in an “intention to treat” analysis of the trial results. However, we found that the intake of nutrients influences breast cancer incidence and tumor phenotype. In particular, higher protein intake and body weight were associated with an increased risk of estrogen receptor (ER) positive breast cancer; and higher CHO intake, primarily starch, was associated with a lower risk of ER positive breast cancer. In addition, higher dietary fat intake was associated with a lower risk of ER negative breast cancer; and higher CHO intake, primarily total sugars, was associated with increased risk of ER negative breast cancer.

These findings are consistent with the experience of native Japanese women who have lower fat and higher carbohydrate (CHO) intake, and have breast cancers that are more often ER negative than Caucasian women in the West (9;10). In addition, as there are major differences in the specific foods consumed in Asia and North America, dietary effects may be much broader than just the distribution of macronutrients and it may be important to look more specifically at different types of food to understand the impact of diet on breast cancer
risk. The types of foods, especially those containing CHO, consumed in countries where breast cancer is less common are substantially different than those found in the Canadian diet.

The purpose of this project is to further investigate the association of CHO intake with breast cancer risk and phenotype in the trial by examining the CHO quality as described by glycemic index (GI). Diet GI was measured in multiple food records collected prior to the diagnosis of breast cancer in a case control study nested within the trial cohort. The goal of this work is to identify if particular types of CHO foods are associated with breast cancer and allow the translation of this information into practical advice.

1.2 Outline of Thesis
The thesis contains four additional chapters:

Chapter 2 Literature review provides a brief description of normal breast tissue, different types of breast cancer and risk factors for breast cancer; as well as an overview of the GI, epidemiological studies of GI and breast cancer risk and the potential biological relationship between GI and the development of breast cancer.

Chapter 3 Methods describes the randomized trial which is the source of subjects and data for this case control study, the procedures used to develop a GI database and to determine diet GI values for multiple food records, and the statistical methods used to analyze the relationship between diet GI values and breast cancer risk and phenotype. This chapter also includes a description and findings of the pilot study we conducted to evaluate our methodology.

Chapter 4 Results includes baseline characteristics of subjects, description of GI values and associations of GI with breast cancer risk overall, and by hormone receptor status in stratified and case-only analyses. This chapter also includes associations of GI with tumor size and nodal status.

Chapter 5 Discussion compares our results with previous studies, describes strengths and limitations of the study and provides a conclusion and direction for future work.
Chapter 2 - Literature Review

2.1 Overview of Breast Cancer Development

Breast anatomy and the stages of breast tissue development provide a foundation for understanding the types of breast cancer that occur and the hormonal factors that influence breast cancer cell growth.

2.1.1 Normal Breast Tissue

Human breasts are composed of parenchymal tissues consisting of a branching ductal system radiating from the nipple to 15 to 20 lobes embedded in stroma of fibrous connective tissue, fat, blood vessels, lymphatics and nerves. Each duct drains a lobe made up of 20 to 40 lobules and each lobule consists of a grape-like cluster of 10 to 100 alveoli which produce milk during lactation. The ducts, lobules and alveoli are lined with layers of epithelial cells. The proportions of fat, fibrous and parenchymal tissue vary greatly between individuals and with menopausal status, weight, number of live births and genetic factors.

Rudimentary breast development begins in utero and then the anatomy undergoes distinct changes at puberty, during menstruation, with pregnancy and lactation and finally at menopause. At birth, a female infant has nipples and a rudimentary ductal system. At puberty the pituitary gland releases follicle stimulating hormone (FSH) and luteinizing hormone (LH) which cause the ovarian follicles to mature and secrete estrogens. The estrogens, primarily 17-estradiol, stimulate the growth and development of the breasts. It takes one to two years after menarche before the ovarian follicles are fully matured and begin to ovulate and produce progesterone. Estrogen and progesterone together contribute to the full development of the ducts, lobules and alveoli. Animal models suggest that the action of estrogen and progesterone on mammary development requires other hormones and growth factors including pituitary growth hormone (GH) and GH induced IGF-1.

Fluctuations in estrogen and progesterone levels during a normal menstrual cycle influence breast morphology. During the first half of the menstrual cycle (follicular phase), under the influence of FSH and LH, estrogen levels increase and peak halfway through the cycle. Ovulation occurs and then a second peak of estrogen occurs in the second half of the cycle (luteal phase) when progesterone levels peak. Estrogen and progesterone induce breast
cell proliferation causing mammary ducts to dilate and alveolar epithelial cells to
differentiate into secretory cells. Most studies have shown that mitotic activity is highest
during the luteal phase of the cycle (14-16).

Estrogen and progesterone act on target tissues, such as the breast, by binding to
intracellular or membrane-bound receptors. An estrogen receptor (ER) is a protein molecule
that contains a specific site to which only estrogens (or closely related molecules) can bind.
There are two different forms of ER, ERα and ERβ. ERα plays a major role in cell
proliferation and growth. The function of ERβ is unclear (17). Progesterone also exerts its
influence on cell growth of by binding to receptors, known as progesterone receptors (PR).
Therefore, when these hormones circulate in the bloodstream, they only exert effects on cells
that contain their receptors.

Estrogen and progesterone also exert a significant effect on ductal, lobular and
alveolar growth during pregnancy. After parturition, many other hormones such as
prolactin, insulin and cortisol play a vital role in milk production.

Ovarian function declines as women approach menopause and through menopause
leading to changes in the epithelial structures and stroma. Although the duct system remains
in the post-menopausal breast, the lobules atrophy, the fibrous tissue decreases and fat
deposition increases.

2.1.2 Development of Breast Cancer

Breast cancer arises as a result of uncontrolled growth of the epithelial cells at the
junction of the terminal duct - lobular unit. It has been estimated that most breast cancers
need about 5 – 10 years to develop from a single malignant cell to a tumor of 5 to 10 mm
diameter (18). There are many histological types of breast cancer but ductal carcinoma is the
most common form and accounts for 85% of breast cancer cases; and lobular breast cancer is
found in only about 15% of cases. Breast cancer is further subcategorized on the basis of
microscopic features as noninvasive (in situ) or invasive (infiltrating). The breast cancer
cases in my masters’ project all have invasive breast cancer.
2.1.3 Phenotype

There are many molecular sub-types of invasive breast cancer but classification according to hormone receptor status has the most important clinical role. The number of estrogen and progesterone receptors found in breast cancer cells after a lumpectomy or mastectomy is described as ER and PR status. ER and PR are usually measured using immunohistochemistry and the cut-off to define positive status varies widely between laboratories as anywhere between 1% and 20% of cells showing the presence of receptors (19). It is estimated that about 70% of tumors are ER positive and about 60 to 65% are PR positive (19). Hormone receptor expression is used to classify breast cancer tumors into four sub-types: ER+/PR+, ER+/PR-, ER-/PR+ and ER-/PR-. ER and PR status of a breast cancer tumor is the most important factor currently available to predict response to treatment. Patients with invasive breast cancer whose tumors are lacking ER and PR receptors do not respond to hormonal based treatment. The first sign of metastatic disease for ER positive tumours is often in bone, whereas ER negative tumors tend to spread to visceral and soft tissue. ER negative breast cancers often recur sooner than ER positive breast cancers; usually within the first five years after diagnosis. The association of risk factors for breast cancer also varies between ER+ and ER- tumours (see section 2.28). Differences in the natural history of hormone receptor positive and negative tumors suggest that they represent distinct sub-type of breast cancer. Their etiology is unknown and it is unclear whether they have separate etiology or if ER negative tumors develop from ER positive tumors.

Staging of tumours describes the extent of the disease which profoundly affects prognosis. The most widely used staging system for breast cancer is that of the American Joint Committee on Cancer (AJCC) which is jointly sponsored by the American Cancer Society and the American College of Surgeons. The AJCC staging system includes both clinical and pathologic prognostic factors that are associated with survival, disease free survival and/or local control. These factors are tumor characteristics known as TNM and they are measured at the time of surgery. T refers to tumor size, N refers to regional lymph nodes and M refers to metastases. The extent of auxiliary lymph node involvement is the major prognostic indicator for later systemic disease as it is evidence of actual metastases growing in regional lymph nodes. Tumor size is the second factor that predicts outcome from disease. The pathology reports for the breast cancer cases in my project describe the ER/PR status,
tumour size and nodal involvement and we will examine the association of these tumour characteristics with dietary factors.

2.2 Risk Factors for Breast Cancer

2.2.1 Overview of Risk Factors for Breast Cancer

Increasing age is the strongest risk factor for the disease. Other important risk factors include: extensive mammographic density, benign breast disease (particularly atypical hyperplasia), family history of breast cancer in a first degree relative, carriage of a known pre-disposing genetic mutation, reproductive events such as age at menarche, parity, age at first birth, breast feeding, age at menopause and use of postmenopausal hormones (HRT), anthropometric factors such as BMI and height, physical activity and dietary factors such as alcohol, fat, and carbohydrate consumption.

Identifying risk factors for breast cancer has enhanced our understanding of the development of the disease, elucidated the international variations in incidence rates, and assisted in predicting the occurrence of breast cancer in individuals. Some risk factors are non-modifiable, such as age, family history, reproductive history and biopsy history. But it may be possible to alter exposure to some risk factors such as mammographic density, body weight and dietary habits, and ultimately reduce the incidence of breast cancer. Therefore, the focus of my thesis is to examine the relationship between a potentially modifiable risk factor, carbohydrate intake and specifically, the glycemic index, and breast cancer risk and phenotype. The literature review is limited to risk factors that are taken into consideration in the statistical analysis of study results.

2.2.2 Age

Breast cancer is most common after menopause as 80% of breast cancers in Canada are diagnosed in women over age 50 (2). The age specific incidence curves for countries with different levels of risk are shown in Figure 2.1. In Canada and Shanghai, the slopes of the curves are very similar for young women and show a rapid increase in breast cancer incidence before menopause, and then the curves diverge. In low risk countries, such as China, the incidence rate slows down considerably and remains relatively stable. However,
in high risk countries, such as Canada, the incidence rate continues to increase, albeit at a much slower rate than in premenopausal years (3).

![Figure 2.1 Age-specific incidence of breast cancer in Canada and Shanghai (2002)](image)

2.2.3 **Mammographic Density**

The radiographic appearance of the breast varies among individuals due to differences in the relative amounts of fat, connective tissue and epithelial tissue. These diverse types of tissues attenuate X-rays to different degrees. Fat is radiolucent and appears dark on a mammogram, while connective tissue and epithelial tissues are radiodense and appear light, an appearance that is known as mammographic density. The breast images in Figure 2.2 illustrate increasing levels of mammographic density from highly radiolucent to almost entirely opaque. Mammographic density can be measured quantitatively using a computer-assisted method that measures the area of dense breast tissue relative to the total breast area as seen on a mammogram (20).
Mammographic density is a strong risk factor for breast cancer; higher levels are associated with higher risk of breast cancer (21). The risk associated with mammographic density is greater than that associated with all other risk factors for breast cancer except age and having a BRCA1 or BRCA2 mutation. A review of ten cohort studies found that women with the most extensive mammographic density have 2.9 to 6 times the risk of breast cancer compared to women with the lowest level of mammographic density (12). Heritability accounts for about 60% of the variation in mammographic density and menopausal status, weight and number of live births account for 20-30% (22). The subjects for my project had extensive mammographic density as this was one of the eligibility criteria for the Diet and Breast Cancer Prevention Trial (DBCPT).

2.2.4 Family History and Genetic Factors

Most women who develop breast cancer do not have a family history of the disease. However, having a first degree relative with breast cancer increases the risk of the disease by a factor of two or three (23). Some genetic mutations, particularly BRCA1 and BRCA2 result in a very high lifetime risk of breast cancer (55-80% lifetime risk) (24). The
prevalence of these mutations is rare and population based studies suggest that mutations in BRCA1 and BRCA2, only account for 2% to 5% of total cases. (IARC world report 2008). Genome wide and candidate gene association studies have identified several common single nucleotide polymorphisms (SNPs) that are associated with small gradients of breast cancer risk, but these SNPs provide only modest improvements in models used to predict breast cancer risk (25).

2.2.5 Reproductive History

The reproductive factors that influence breast cancer risk can be divided into three categories: menstrual (age at menarche and age at menopause), pregnancy (parity, age at first birth and duration of breast feeding) and hormonal factors (use of oral contraceptives and HRT). Reproductive factors such as early menarche, late menopause, late (over 30) age at first birth and nulliparity, may reflect increased lifelong exposure to endogenous reproductive hormones. The total duration of exposure to endogenous reproductive hormones appears to be very important in contributing to breast cancer risk (24).

2.2.5.1 Menstrual Factors

Epidemiological studies have consistently found that early age at menarche is associated with increased overall risk of breast cancer but the effect is stronger for premenopausal breast cancer. Relative risk for premenopausal breast cancer is reduced by about 7% for each year that menarche is delayed after age 12 years, and by about 3% for postmenopausal breast cancer (26).

Menopause has a protective effect on breast cancer risk as evidenced by the markedly decreased age-specific incidence rate of the disease beginning at age 50. Late age of menopause increases the risk of breast cancer. For each year menopause is delayed, there is about 3% increase in breast cancer risk (27). The effect is similar whether menopause occurs naturally or as a result of bilateral oophorectomy.
2.2.5.2 Pregnancy Factors

Full term pregnancy is followed by a transient increased risk of breast cancer which gradually declines over five to 10 years (28-30). The long term protection of pregnancy is inversely proportional to the age of first birth. Women who had their first birth before age 20 years had 30% lower risk of breast cancer than women with a first birth after age 35 years (28;31). Women who have had at least one full term pregnancy have about a 25-30% reduction in breast cancer risk compared with nulliparous women (32;33). In addition, the level of protection increases with the number of full term pregnancies with a further reduction in risk of 7% for each birth after the first, independent of breast feeding.

Breastfeeding provides additional risk reduction and the magnitude of the effect is contingent on the duration of breastfeeding. According to a meta-analysis from 47 epidemiological studies from 30 countries, the relative risk of breast cancer is reduced by 4% for every 12 months of breast feeding (34).

2.2.5.3 Sex Hormones

Several lines of evidence suggest that cumulative, excessive exposure to estrogen is associated with the development of breast cancer. The age specific breast cancer incidence curves reflect the usual pattern of average endogenous estrogen levels. There is a rapid increase in incidence rate during the pre-menopausal years when endogenous estrogen levels are high and a reduced incidence rate at the time of menopause, when estrogen levels decline rapidly. Several epidemiological studies have shown that endogenous estrogen levels are positively associated with breast cancer risk, but primarily in ER positive tumors (11). Also, the selective estrogen receptor modulator, such as Tamoxifen, which blocks estrogen from receptors in breast tissue has been useful in the treatment of breast cancer and has reduced the incidence of ER positive breast cancer in women at high risk (35).

Exposure to exogenous female hormones such as oral contraceptives (OC) and menopausal hormone replacement therapy (HRT) may also increase risk for breast cancer among current and recent users. OC and HRT are administered for variable but usually extended periods of time and would therefore elevate serum levels of estrogen and progesterone and affect breast tissue growth.
A 1996 analysis of data from 54 international epidemiological studies conducted by the Collaborative Group on Hormonal Factors in Breast Cancer found that women who were current users of OC had an elevated risk of developing breast cancer (RR: 1.24; 95%CI: 1.15-1.33), and for women who stopped using OC, their risk declined progressively over 9 years and after stopping use for 10 years or more there was no increased risk (RR: 1.01; 95% CI: 0.96-1.05). Duration of use, age at first use, dose and type of hormones had little effect (36). However, the Women’s CARE study found that among women from 35 to 64 years of age, current or former OC use was not associated with an increased risk of breast cancer (37). Therefore, the short term risk associated with OC use may be more relevant for younger women.

Several observational studies and a randomized trial, The Women’s Health Initiative (WHI) Estrogen-Progestin Study, have consistently shown that current and recent users of HRT have an increased risk of invasive breast cancer (38;39) (27;40;41). Two large prospective studies (40;42) suggest there is an increased risk among current users of estrogen only therapy and a still higher risk among users of combined HRT. The WHI Estrogen-Progestin Study found that after 5 years of follow up, women taking combined HRT had a 24% increase in breast cancer risk compared with women taking the placebo. However, any excess risk seems to disappear two to five years after estrogen-progestin therapy is discontinued (27;39).

2.2.6 Height

Case-control and cohort studies suggest that adult height is positively associated with breast cancer risk but the effect may vary by menopausal status. The relationship seems less consistent for pre-menopausal than postmenopausal breast cancer (43) (44). The increase in relative risk of postmenopausal breast cancer is estimated to be about 7-10% for each additional five centimeters in height (43;44).

The underlying mechanisms for the association between height and breast cancer remain unclear, but height may be a marker for other exposures that promote growth in childhood and influence breast cell proliferation. Factors that influence both attained height and breast cancer are increased levels of growth hormone, insulin, IGF-1 and sex steroids as well as adequate energy intake during childhood.
2.2.7 **Body Weight**

The association of body weight and breast cancer risk varies by menopausal status. Most studies show that obesity has a protective effect on the risk of premenopausal breast cancer. Estimates from cohort studies suggest a 7–14% decreased risk of premenopausal breast cancer per 5 kg/m² (43;45;46). Epidemiological studies consistently demonstrate that overweight and obesity, increase the risk of postmenopausal breast cancer (45;47). A meta-analysis of 34 cohort studies indicated that every 5 kg/m² increase in body mass index (BMI) was associated with a relative risk of post-menopausal breast cancer of 1.12 (95% CI 1.08-1.16, p <0.0001) (45). However, three large cohort studies: WHI, Nurses’ Health Study and EPIC, have failed to find this relationship among post-menopausal women who used HRT (44;48;49). The relationship of weight gain to postmenopausal breast cancer is similar to that of BMI. Several prospective studies have shown that large weight gain during adulthood (after age 18) is positively associated with postmenopausal breast cancer in HRT nonusers, and losing weight after menopause can reduce risk (50;51).

Abdominal fatness, as measured by waist circumference or waist to hip ratio (WHR) may also affect breast cancer risk. There have been a small number of studies investigating abdominal fatness and premenopausal breast cancer and overall they have failed to find a significant association with either waist circumference or WHR. Increased central adiposity may impact risk of postmenopausal breast cancer but the results from studies on the effect of abdominal fat are much less consistent than studies on weight gain and BMI (46).

One biologically plausible explanation for the association between obesity and post menopausal breast cancer risk is related to an increase in the serum concentration of free estradiol. The concentration of free estradiol increases in obese women after menopause due to an increase in the production of estrogens from the conversion of androgens by the aromatase enzyme in adipose tissue and a decrease in the serum concentration of sex hormone binding globulin (SHBG). After menopause when the ovaries stop producing estrogen, adipose tissue becomes the primary source of endogenous estrogen so overweight and obese women are exposed to higher levels of estrogen than lean women. There is substantial evidence that high concentrations of endogenous estrogen are associated with increased risk of postmenopausal breast cancer (44). The mechanism underlying the inverse association of obesity with risk of pre-menopausal breast cancer is not clear. Obesity may be
associated with more anovulatory menstrual cycles which would reduce exposure to endogenous sex hormones and lower breast cancer risk (52).

2.2.8 Risk factors by Estrogen Receptor Status

There is considerable heterogeneity in the risk factors for breast cancer by receptor status. There are marked differences in the effect of age on breast cancer risk by estrogen receptor status. Yasui and Potter demonstrated that the age specific incidence rates are similar to those shown in figure 2.1. The age specific incidence rates for ER positive tumours are similar to those seen in Canada, USA and Europe with most occurring after menopause. The age specific incidence rate for ER negative breast cancer is like the Japan and other Asian countries where it peaks before menopause and then the incidence rate remains relatively stable (53). Within the last ten years many large observational studies have examined the effect of other risk factors on breast cancer by ER and PR status (54). In general, risk factors that influence endogenous levels of estrogen are more strongly associated with ER positive, than ER negative tumours. There is consistent evidence among epidemiological studies that parity, age at first birth, BMI and weight gain are associated with ER positive breast cancer. The evidence is less clear for the effect of age at menarche and lactation. The risk of breast cancer associated with benign breast disease and family history do not seem to be related to ER status. There are racial differences in the incidence of hormone receptor negative breast cancer. ER negative/PR negative breast cancer appears to be more common among African Americans than Whites, Asians and Latinas (55-57). These findings contribute to our understanding of the etiology of breast cancer and reinforce the importance of considering hormone receptor status when assessing the impact of exposures on risk.

2.2.9 Dietary Factors

The wide international variation in breast cancer incidence and the large increases in rates among the offspring of migrants from low – to high-risk countries suggest that environmental factors play a role in the development of breast cancer. Many foods, nutrients and phytochemicals have been investigated as possible environmental determinants of breast cancer. Research about nutritional factors often begins with a review of ecological data to
compare the international variations in per capita consumption of different foods or nutrients in relation to breast cancer incidence and/or mortality. Any significant associations between nutritional factors and cancer risk provide motivation for animal studies to test their effect in vivo. Results from ecological and laboratory research generate specific theories that are examined using epidemiological studies, starting with case-control and cohort studies involving large numbers of subjects. These observational studies describe eating habits of a population more accurately than ecological studies through the use of food frequency questionnaires (FFQ) and 24 hour food recalls. Therefore, these studies provide more information about the association of nutrition with cancer incidence but clinical trials are required to demonstrate a direct effect of a dietary factor on risk.

Research about a role for dietary factors in breast cancer risk has examined a wide variety of foods from every food group as well as tea, coffee and sugar; and numerous macronutrients, vitamins, minerals, phytochemicals and carcinogens found in foods and produced during cooking. I will limit my review to alcohol, as it is the only dietary factor for which there has been consistent and convincing evidence of an effect on breast cancer risk; and fat, and carbohydrate because they were the basis of the intervention trial that prompted this master’s thesis.

2.2.9.1 Alcohol

There has been one international correlation study of alcohol consumption and breast cancer risk and mortality. The per capita alcohol consumption from 46 countries was positively correlated with breast cancer incidence \((r = 0.65)\) and mortality \((r = 0.31)\) until the regression models were adjusted for fat consumption; and then the association disappeared \((58)\). These results are limited by their estimate of alcohol intake because it is based on per capita consumption levels which may not reflect the alcohol consumed by women. There is substantial evidence from observational studies for a significant association between alcohol consumption and an increased risk of breast cancer. A meta-analysis of 53 epidemiological studies published in 2002 by the Collaborative Group on Hormonal Factors in Breast Cancer, found that compared to women who reported drinking no alcohol, women who drank \(35 – 44\) g alcohol per day had a relative risk of breast cancer of \(1.32\) \((1.19-1.45, P<0.00001)\), and women who drank \(\geq 45\) g per day alcohol had a relative risk of \(1.46\) \((1.33 – 1.61,\)
P<0.00001). There is a dose response effect as the relative risk of breast cancer increased by 7.1% (5.5-8.7, P<0.0001) for each additional 10 g per day intake of alcohol (59). Similar results were reported recently by the National Institutes of Health –AARP Diet Study (60) and the Women’s Health Study (61). Further examination by tumour receptor status suggests that the increased risk due to alcohol may be limited to ER positive tumours (54;62) (61) (63). This relationship may be due to the stimulatory effects of ethanol on the production of endogenous estrogen, expression of ER and/or proliferation of ER positive breast cancer cells (63).

2.2.9.2 Total Fat

The hypothesis that increased dietary fat increases the incidence of breast cancer stems from the results of human ecological studies and animal experiments. International studies show a striking positive correlation between breast cancer rates and estimates of fat consumption (Figure 2.3) (64;65).

![Figure 2.3](image-url) Correlation of percent energy from fat with breast cancer incidence
Adapted with permission from Ref. 64; ©1987 Oxford University Press. All rights reserved.

Animal studies show some evidence of promotional effects of dietary fat on the development of mammary tumors but it is unclear whether the association is independent of energy intake (8;66).
Epidemiological evidence from cohort and case control studies has however been inconsistent. A meta-analysis of 31 case control and 14 cohort studies published in 2003 showed a modest, yet significant positive association for intakes of total fat (RR for all studies: 1.13; 95% CI: 1.03 – 1.25) (67). The summary relative risks for case control studies and cohort studies both showed positive, but not statistically significant associations. The meta-analysis also reported a positive association for saturated fat with breast cancer risk (RR: 1.17; 95%CI: 1.06-1.29). Three cohort studies have published their results since the meta-analysis. The Malmo Diet and Cancer cohort of postmenopausal women found that total fat (p = 0.03), and polyunsaturated fat (p = 0.0009) were positively associated with breast cancer risk (68). The National Institutes of Health-AARP Diet and Health Study also found a positive association for total fat (RR 1.15 95% CI: 1.05-1.26) and all subtypes of fat with breast cancer risk in post-menopausal women (69). The European Prospective Investigation into Cancer and Nutrition (EPIC), a cohort of 319,826 pre- and postmenopausal women from 10 European countries, found a significant association between saturated fat and breast cancer risk but failed to find an association with total fat intake (70). In addition, data from the cohort of controls from the Women’s Health Initiative Diet Modification Trial found that total fat intake was positively associated with breast cancer risk when measured with 4-day food records, but not with food frequency questionnaires (71).

Two randomized clinical trials have been conducted to determine if a low fat, high carbohydrate diet would reduce breast cancer risk: The Women’s Health Initiative Diet Modification Trial and the Diet and Breast Cancer Prevention Study. They are described in section 2.2.9.5.

2.2.9.3 Carbohydrate

An association between carbohydrates and breast cancer development was reported in 1975 when Hems and Stuart correlated the mortality rate from breast cancer of single, post menopausal women in 10 North American and European countries with per capita consumption of sugar and starch. They found a positive association for sugar (r = 0.84) and a negative association for starch (r= -0.71) (Figure 2.4) (72). These observations were confirmed in another ecological study of 27 countries (7;73) and provide support for further investigation of the relationship of different types of carbohydrate and breast cancer.
There have been four experiments on the effects of dietary carbohydrates on the promotion of mammary carcinogenesis in laboratory animals given a carcinogen to induce mammary tumors. Hoehn and Carroll found that rats fed dextrose or sucrose developed significantly more mammary tumors than those fed wheat, rice or potato starch (74). Gridley et al used mixed diets but changed the source of carbohydrate from dextrin to sucrose. They found that mice fed sucrose had more tumors than mice fed a comparable diet containing dextrin (75). Klurfeld et al found that rats fed lactose had fewer tumors than those fed either sucrose or starch; and that consumption of starch was positively associated with tumor weight (76). Thompson et al fed rats five different diets of identical macronutrient composition but cornstarch was replaced with increasing amounts of beans (77). They observed that bean consumption was related to a dose-dependent reduction in mammary cancer incidence and tumor burden. These animal studies demonstrate that dietary carbohydrates have varied and significant effects on the development of mammary tumors.

The relationship between total carbohydrate intake and breast cancer risk was considered in 10 cohort studies (78-87). No studies found an association of total CHO intake with overall risk of breast cancer, alone or after adjustment for risk factors, energy intake and BMI. Five studies found associations with sub-groups of subjects. Four studies found a positive association of total CHO with subjects who: had localized, Grade I tumours (82), ER positive/PR negative tumors (79), ER negative breast cancer (81) or were premenopausal
or less than 50 years old (80). One study found total CHO intake was negatively associated with breast cancer risk in premenopausal women with a BMI < 25 kg/m² (87).

A protective role for dietary fibre against breast cancer has been hypothesized because fibre intake has been shown to influence estrogen metabolism. High fibre intakes reduce levels of serum estrogen by inhibiting intestinal re-absorption of estrogen and reducing estrogen synthesis by inhibiting the activity of aromatase (human estrogen synthetase). However, most prospective studies (87;88) (89-94), but not all (95) (96) have failed to find a protective effect of dietary fibre with overall breast cancer risk. Giles et al found different effects when data from the Melbourne Collaborative Cohort Study was analyzed by hormone receptor status. They found total fibre was positively associated with ER positive/PR positive breast cancer (RR 1.36, 95% CI: 1.10-1.67), and negatively associated with ER negative/PR negative breast cancer (RR 0.65, 95% CI 0.43-0.99) (82).

The UK Women’s Cohort Study found an inverse relationship between dietary fibre and breast cancer in pre-menopausal (p trend = 0.01), but not in post-menopausal women (p trend = 0.97) (94).

Epidemiological studies about GI and breast cancer risk are described in section 2.2.10.2.

2.2.9.4 Limitations of Observational Studies

The lack of convincing evidence from observational epidemiological studies of an association between diet and breast cancer risk may be due, at least in part, to the populations studied and the limitations of the instruments used to measure dietary intake. Observational studies have examined populations with a narrow range of fat and carbohydrate intakes which make it difficult to detect small to moderate associations with disease using diet assessment instruments known to have significant measurement error. Prospective studies usually rely on food frequency questionnaires (FFQ) to estimate usual intakes of foods and nutrients rather than 24 hour recalls or food records. FFQs are more efficient, economical and reflect dietary habits over longer time periods than 24 hour recalls and food records. However, compared to food records, FFQ are known to contain more measurement error (97;98). Food records and recalls provide detailed information about the type and amount of foods consumed but they capture short periods of time and may be subject to within subject
variation. This limitation can be ameliorated by collecting multiple food record days. Collecting 7 to 14 record days has been suggested to describe the intake of most nutrients and food groups (99). Three prospective studies of diet and prospective breast cancer risk have collected dietary data using both a FFQ and food records and they all found a positive association with fat (71), or saturated fat, with food records but not FFQs (68;100). Therefore, using FFQs may have attenuated the associations between dietary fat and breast cancer.

Epidemiological studies within homogeneous populations have been hampered by a relatively narrow range of fat intake where the lowest level intake was about 30% of calories from fat, which is above that found in low and medium risk countries. Therefore, controlled clinical trials would be required to examine a wider range of fat intake and determine if changing dietary fat intake would reduce breast cancer risk.

2.2.9.5 Clinical Trials of Low Fat, High Carbohydrate Diets

Two randomized, primary prevention trials have been conducted to determine the effect of a low fat, high carbohydrate diet on breast cancer risk: Women’s Health Initiative (WHI) Randomized Controlled Dietary Modification Trial and the Diet and Breast Cancer Prevention Study (DBCPT). The WHI began in the USA in 1993 and the DBCPT began in Canada in 1988 and both studies ended in 2005.

The WHI Dietary Modification trial recruited 48,835 healthy post-menopausal women and randomly assigned 40% to the dietary modification intervention group and 60% to the control group. The women in the intervention group were asked to reduce their fat intake to 20% of energy from fat, and increase consumption of fruit, vegetables and grains. Controls were not asked to make any dietary changes. After mean follow up of 8.1 years, there was no difference in overall risk of invasive breast cancer between the two groups (hazard ratio 0.91, 95% CI: 0.83-1.01). However, secondary analysis suggested significant risk reduction among women with a high fat diet at baseline (hazard ratio 0.78, 95% CI: 0.64-0.95). When the analysis was stratified by hormone receptor status, they found the intervention was associated with a lower risk of ER positive/ER negative breast cancer (hazard ratio 0.64, 95% CI: 0.49-0.84). There was no association of diet with tumour size, grade, stage or nodal status (101).
Nutritional assessment in the WHI was based on data from baseline and annual FFQs. However, all subjects also completed 4-day food records prior to randomization. The results were analyzed from a nested case-control study of subjects from the control arm of the WHI using food records and then FFQ. They found total fat was associated with an increased risk for breast cancer (RR 1.82, 95% CI: 1.12-2.98, p trend = 0.02) for food records, but not for FFQ (RR 0.67, 95% CI: 0.33-1.37, p trend = 0.24) (71).

The DBCPT recruited 4,690 women between 30 and 65 years of age who were at increased risk for breast cancer because they had extensive breast density. The subjects were also randomly assigned to either an intervention or control group. The women in the intervention group were asked to reduce their fat intake to 15% of energy from fat and increase their CHO intake in order to maintain body weight. Women in the control group were not asked to change their diet. All subjects were asked to provide 3-day food records at baseline and at every clinic visit throughout the trial follow up. After an average of 10 years of follow up, there was no significant difference between the groups in the risk of developing invasive breast cancer (hazard ratio 1.19, 95% CI: 0.91-1.55) (unpublished). The recruitment process, dietary intervention and data collection for the trial are described in the methods 3.1.

In a case-control study nested within the cohort of DBCPT participants, we found that weight, fat, protein and carbohydrate were significantly associated with risk of invasive breast cancer. Figure 2.5 shows the results of the logistic regression analysis for the association of weight and these nutrients with breast cancer risk overall, and by ER status. The analysis was conducted with continuous data collected during post-randomization but the risk estimates are presented as odds ratio across the inter-quartile range.
Breast cancer risk overall was positively associated with total protein intake. ER positive breast cancer risk was positively associated with body weight and total protein intake, and negatively associated with starch. ER negative breast cancer risk was positively associated with intake of total CHO and total sugars, and negatively associated with total fat intake. In the same analysis using baseline food records, the only significant association we observed was an inverse association between starch and risk of ER positive breast cancer risk.

We also examined the associations of weight and nutrient intake with tumor size and nodal status. Tumor size was positively associated with baseline and post-randomization body weight (p = 0.02). Nodal status was positively associated with post-randomization fat intake (p = 0.01), and negatively associated with total CHO intake (p = 0.02).

These findings related to total CHO, starch and total sugars were the motivation to further explore the relationship between breast cancer and carbohydrate by examining the effect of the glycemic index on breast cancer risk and phenotype.
2.2.10 **Glycemic Index**

Glycemic index (GI) is a relative measure for ranking carbohydrate containing foods based on their impact on blood glucose levels. Different foods containing equal amounts of carbohydrate can produce a wide range of glycemic responses (increases in blood glucose and insulin) due to variations in the type(s) of carbohydrate and starch they contain. GI is defined as the incremental area under the blood glucose response curve following the intake of 50 g available carbohydrate from food compared with the glucose area generated from a similar amount of white bread or glucose (102). When compared to low GI foods, the consumption of food with a high GI results in a higher increase in blood glucose levels, leading to a higher increase in blood insulin levels (103).

The GI of a food is largely determined by the type of sugar, nature of the starch, cooking and processing. Each mono and disaccharide has a different GI value, ranging from about 20 – 25 for fructose and lactose, 60 for sucrose and 100 for glucose. Therefore, their relative contribution to the CHO content of a food has a strong influence of the GI. The effect of starch on GI varies with the ratio of amylose to amylopectin, and the particle size and degree of gelatinization after processing and cooking. Amylose which is composed of long chains of glucose units is more resistant to hydrolysis than the branched formation of amylopectin, and therefore starchy foods with a higher proportion of amylose will have a lower GI. Grinding, rolling and milling starchy foods, like breakfast cereals reduces particle size which increases the rate of digestion and absorption. Similarly, cooking with moist heat gelatinizes starch, expanding the granules which enhances the hydrolysis of glucose units and hastens the glycemic response. The effect of dietary fibre on GI varies with the type of fibre. Insoluble fibre has little effect whereas large amounts of viscous soluble fibre will lower GI. The effects of fat and protein on glycemic response operate independently of CHO. This is evident as the glycemic response to mixed meals can consistently be predicted from the formula based on amount of available CHO and the measured GI of each food consumed (104-106).

Based on the calculation for the GI of mixed meals, the overall diet GI is defined as the weighted mean of the GI of individual CHO containing foods and is calculated as:

\[ \sum (\text{available CHO in g for each food} \times \text{GI for each food}) / \text{total available CHO in g/day} \] (107).
Glycemic load is a measure of the combined effect of CHO quantity and type \((108)\) and is calculated as: \(\sum(\text{available CHO in g x GI/100 for each food})\). The glycemic load is highly correlated with CHO intake \((r = 0.9)\) \((109-111)\) and largely reflects the amount of CHO consumed.

2.2.10.1 Nutrients, Foods and Diet GI

Diet GI is moderately correlated with the intake of many foods and nutrients. It tends to be inversely associated with mono- and disaccharides, fibre, protein, vegetables, fruit and dairy products, and positively associated with starch, fat and wheat bread \((107;109)\). These relationships are consistent with the GI values of the foods. Fruits, vegetables and dairy products which are rich in simple sugars, are low GI foods. Some starchy foods like pasta, oatmeal and legumes have a low GI, but most other breads, cereals and potatoes tend to have a high GI value. The ability of the diet GI to reflect the carbohydrate composition of the diet supports its usefulness as a measure of glycemic response, but the diet GI also has independent effects that cannot be fully explained by the intake of nutrients and combination of foods. Therefore, the diet GI could provide additional information about dietary exposures and their associations with breast cancer risk and phenotype.

2.2.10.2 Glycemic Index, Glycemic Load and Breast Cancer Risk

There have been 17 observational studies examining the association of GI and glycemic load with breast cancer risk and phenotype, 13 cohort and 4 case-control studies \((78-80)\) \((81-87;112-118)\) (Table 2.1). There have been no randomized clinical trials. There have been nine studies conducted in North America, six in Europe, one in Asia and one in Australia. Three studies found GI was positively associated with overall breast cancer risk \((86;112;118)\) and six studies found a positive association with subgroups only \((79)\) \((81;83-85;113)\).

In August 2008, Mulholland et al published a systematic review and meta-analysis of 14 epidemiological studies of GI and breast cancer risk by menopausal status \((119)\). They considered all observational studies, 10 cohort and 4 case-control studies, published at the time in a manner consistent with usual practice for meta-analysis. The case-control studies were excluded because of significant heterogeneity in their results. They used cohort studies
that categorized diet GI or Glycemic load into quartiles (81) or quintiles (87), (84-86;112-114) in order to determine relative risk estimates for the highest versus the lowest category of diet GI and Glycemic load. Therefore, two cohort studies were excluded from the meta-analysis because they only examined GI as a continuous, and not a categorical variable (82;83). Neither of these studies found an association with GI and overall risk of breast cancer. The meta-analysis did not include three cohort studies published in 2009 and 2010. Neither study found an association with GI and overall breast cancer risk (79;80).

When they combined the results for the remaining eight cohort studies, they found a non-significant positive association between the highest versus the lowest category of GI and pre-menopausal (RR: 1.14; 95% CI: 0.95-1.38) and post-menopausal (RR: 1.11; 95% CI: 0.99-1.25) breast cancer risk. However, when they considered the quality of the nutrient data, and limited their analysis to studies that used a food frequency questionnaire (FFQ) with ≥100 items, a significant positive association emerged for both premenopausal (RR: 1.20; 95% CI: 1.01-1.43) (86;87;112;113) and post menopausal (RR: 1.10; 95% CI: 1.02-1.19) breast cancer (81;112;113).
Figure 2.6  Meta analysis of glycemic index and breast cancer risk
* studies with ≥100 items on FFQ
Reprinted with permission from Ref 119; © 2008 Nature Publishing Group.
All rights reserved.
<table>
<thead>
<tr>
<th>Lead Author (date)</th>
<th>Country</th>
<th>Study&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Mean Follow-up</th>
<th>No. of Cases</th>
<th>Cohort Size</th>
<th>No. of items on FFQ&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Median GI (IQR)&lt;sup&gt;c&lt;/sup&gt;</th>
<th>RR&lt;sup&gt;d&lt;/sup&gt;</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linos (2010)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>USA</td>
<td>Nurses’ Health II</td>
<td>8 years</td>
<td>455</td>
<td>39,268</td>
<td>124</td>
<td>55 (52-58)</td>
<td>1.18</td>
<td>0.88-1.58</td>
</tr>
<tr>
<td>Larsson (2009)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Sweden</td>
<td>Swedish Mammography Cohort</td>
<td>17.4 years</td>
<td>2,952</td>
<td>61,433</td>
<td>67 &amp; 96</td>
<td>57 (52-61)</td>
<td>1.08</td>
<td>0.96-1.21</td>
</tr>
<tr>
<td>Wen (2009)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>China</td>
<td>Shanghai Women’s Health</td>
<td>7.4 years</td>
<td>616</td>
<td>74,942</td>
<td>77</td>
<td>71 (64-77)</td>
<td>1.03</td>
<td>0.79-1.34</td>
</tr>
<tr>
<td>Lajous (2008)</td>
<td>France</td>
<td>EPIC (France)</td>
<td>9 years</td>
<td>1,812</td>
<td>62,739</td>
<td>208</td>
<td>55 (44-66)</td>
<td>1.14</td>
<td>0.99-1.32</td>
</tr>
<tr>
<td>Sieri (2007)</td>
<td>Italy</td>
<td>Hormones and Diet in the Etiology of Breast Tumors (ORDET)</td>
<td>11.5 years</td>
<td>289</td>
<td>8,959</td>
<td>107</td>
<td>56 (52-59)</td>
<td>1.57</td>
<td>1.04-2.36</td>
</tr>
<tr>
<td>Giles (2006)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Australia</td>
<td>Melbourne Collaborative Cohort</td>
<td>9.1 years</td>
<td>324</td>
<td>12,273</td>
<td>121</td>
<td>49 (46-53)</td>
<td>0.98</td>
<td>0.88-1.10</td>
</tr>
<tr>
<td>Nielsen (2005)&lt;sup&gt;cf&lt;/sup&gt;</td>
<td>Denmark</td>
<td>Diet, Cancer and Health Cohort (post)</td>
<td>6.6 years</td>
<td>634</td>
<td>23,870</td>
<td>192</td>
<td>64</td>
<td>0.94</td>
<td>0.80-1.10</td>
</tr>
<tr>
<td>Silvera (2005)</td>
<td>Canada</td>
<td>National Breast Screening</td>
<td>16.6 years</td>
<td>2,518</td>
<td>49,111</td>
<td>86</td>
<td>55 (43-68)</td>
<td>0.88</td>
<td>0.63-1.22</td>
</tr>
<tr>
<td>Higginbotham (2004)</td>
<td>USA</td>
<td>Women’s Health</td>
<td>6.8 years</td>
<td>946</td>
<td>38,446</td>
<td>131</td>
<td>53 (50-55)</td>
<td>1.03</td>
<td>0.84-1.28</td>
</tr>
<tr>
<td>Holmes (2004)</td>
<td>USA</td>
<td>Nurses’ Health</td>
<td>18 years</td>
<td>4,092</td>
<td>88,678</td>
<td>61+</td>
<td>53 (49-58)</td>
<td>1.02 (pre)</td>
<td>0.82-1.28</td>
</tr>
<tr>
<td>Frazier (2004)</td>
<td>USA</td>
<td>Nurses’ Health II</td>
<td>361</td>
<td>47,355</td>
<td>131</td>
<td>56 (53-60)</td>
<td>1.47</td>
<td>1.02-1.30</td>
<td></td>
</tr>
<tr>
<td>Cho (2003)</td>
<td>USA</td>
<td>Nurses’ Health II (pre)</td>
<td>8 years</td>
<td>714</td>
<td>90,655</td>
<td>133 &amp; 142</td>
<td>55 (50-58)</td>
<td>1.05</td>
<td>0.83-1.33</td>
</tr>
<tr>
<td>Jonas (2003)</td>
<td>USA</td>
<td>CPS II Nutrition Cohort (post)</td>
<td>5 years</td>
<td>1,442</td>
<td>63,307</td>
<td>68</td>
<td>53 (46-60)</td>
<td>1.03</td>
<td>0.87-1.22</td>
</tr>
</tbody>
</table>

<sup>a</sup> Study name.

<sup>b</sup> FFQ: Food Frequency Questionnaire.

<sup>c</sup> IQR: Interquartile Range.

<sup>d</sup> RR: Relative Risk.

<sup>e</sup> Year of publication.

<sup>f</sup> Study conducted as part of a larger cohort study.
Legend:

- All cohort studies include pre- and postmenopausal women except where indicated. All studies are prospective except Frazier
- FFQ are self-administered at the start of the study, except Larsson et al administered at baseline and 1997, Holmes et al completed multiple FFQ and Cho et al completed two FFQ
- IQR represents the mean of the lowest and highest categories
- Risk estimate for overall risk of breast cancer from multivariate model
- Study not included in meta-analysis
- Analysis not adjusted for energy intake

<table>
<thead>
<tr>
<th>Lead Author (date)</th>
<th>Country</th>
<th>Studya</th>
<th>Source of Controls</th>
<th>No. Cases</th>
<th>No. Controls</th>
<th>No. of Items on FFQ</th>
<th>Median GI (IQR)</th>
<th>OR</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCann (2007)</td>
<td>USA</td>
<td>WEBb</td>
<td>Population-based</td>
<td>1,166</td>
<td>2,105</td>
<td>113</td>
<td>55 (50-59)</td>
<td>1.02 (pre)</td>
<td>0.68-1.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.80 (post)</td>
<td>0.61-1.03</td>
<td></td>
</tr>
<tr>
<td>Lajous (2005)</td>
<td>Mexico</td>
<td></td>
<td>Population-based</td>
<td>475</td>
<td>1,391</td>
<td>62</td>
<td>0.84</td>
<td>0.62-1.15</td>
<td></td>
</tr>
<tr>
<td>Levi (2002)c</td>
<td>Switzerland</td>
<td></td>
<td>Hospital-based</td>
<td>331</td>
<td>534</td>
<td>79</td>
<td>65 (52-80)</td>
<td>1.25</td>
<td>0.83-1.87</td>
</tr>
<tr>
<td>Augustin (2001)c</td>
<td>Italy</td>
<td></td>
<td>Hospital-based</td>
<td>2,569</td>
<td>2,588</td>
<td>78</td>
<td>53 (50-56)</td>
<td>1.36</td>
<td>1.14-1.64</td>
</tr>
</tbody>
</table>

Legend:

- All studies include pre- and postmenopausal women
- WEB = Western New York Exposure and Breast Cancer Study
- Analysis not adjusted for weight or BMI
Four studies examined the association of the diet GI with breast cancer according to estrogen receptor/progesterone receptor (ER/PR) status and the results were inconsistent (79;81-83). The Swedish study reported that GI was positively associated with ER positive/Progesterone negative receptor tumours (83). The Danish study found a positive association of GI with ER negative tumours (83). Two studies found no association of GI and ER status (81;82).

Glycemic load was positively associated with overall risk of breast cancer in two cohort studies (79;112) and in two case-control studies (116) (118). In three out of four studies, the results were the same (112) (118) or similar (79) to GI. In three studies, glycemic load was positively associated with breast cancer risk in subgroups only (80;81) (82). Two studies were consistent with GI (79;81), but in one study the association was seen in GL only (80). Two studies found total CHO and glycemic load were positively associated with ER negative tumors (79;81).

Generally the findings for total CHO intake were consistent with those for glycemic load, but not with GI. Glycemic load is likely a surrogate for total CHO and largely reflects the amount of carbohydrate consumed, whereas GI offers some unique information about carbohydrate quality.

Although not direct measures of GI, there have been four case control studies of consumption of sugar and desserts and they all showed a positive association with breast cancer risk (120) (121) (122;123). Two studies also measured total carbohydrate intake but neither found an association with breast cancer (121;123).

2.2.10. 3 Limitations related to measurements of Diet GI

All of the studies discussed above measured food intake with a FFQ. As discussed in section 2.2.9.4, FFQ are associated with some measurement error and five studies attempted to validate the macronutrient results from their FFQ against multiple food records or recalls. In these studies, the correlation coefficients for total carbohydrate intake were between 0.47 and 0.64 for the two measures (79;81;83;87;114).

However, no study used food records or recalls to validate their measure of GI. When the FFQs used in these studies were designed, foods were combined without consideration of their GI value. Therefore, if foods with very different GI values are grouped
together, it would be difficult for one GI value to represent accurately all the foods in the group despite the usual weighting procedures. This issue is particularly relevant because each study used a different FFQ of variable length, between 61 to 208 food items. Shorter questionnaires may have compiled more foods into each category, or omitted some major types of carbohydrate foods. Therefore, there was likely substantial error in measuring GI that contributed to the attenuation and inconsistency of the results (124).

2.3 Biologic Plausibility

There are some biological mechanisms that may underlie an association between carbohydrate, GI, and breast cancer. Theories of potential mechanisms have developed around a role for insulin in the etiology of breast cancer as it is influenced by GI and insulin levels are related to conditions and physiological effects that may directly or indirectly affect breast cancer, such as obesity, insulin resistance, type 2 diabetes, hormones, growth factors, serum lipids, and cell proliferation (125-128) (129;130) (Figure 2.7). Low GI diets can reduce postprandial glucose and insulin levels, improve insulin sensitivity, and may reduce risk of type 2 diabetes, all of which may have beneficial effects for breast cancer.

Excess insulin increases levels of free Insulin-like Growth Factor 1 (IGF1), and IGF1 stimulates mitosis and inhibits apoptosis of neoplastic cells (131). IGF1 circulates in the blood bound to insulin-like growth factor binding proteins (IGFBPs) of which IGFBP-3 is the principal carrier form (132). Insulin is secreted by the beta cells of the pancreatic islets in response to the blood glucose concentration, and binds to the insulin receptor to activate signaling pathways concerned with energy metabolism and other functions. Insulin is a structural homologue of IGF1, is required for the production of IGF-1, and is an independent predictor of serum levels of IGF1 (133). Insulin and IGF 1 are potent mitogens to epithelial cells including breast cells and function in an integrated fashion to promote cell growth and survival (134). Epidemiological studies support a role for IGF1 in breast cancer development. A meta-analysis of 17 cohort studies showed that IGF1 was positively associated with breast cancer risk for pre- and postmenopausal women but was limited to ER positive tumours (135). The EPIC study investigated correlations between dietary intake of nutrients and IGF levels. They found that sugar was inversely associated with Insulin-like Growth Factor Binding Protein -1 (IGFBP-1). Although present in smaller concentrations than other
binding proteins, IGFBP-1 can influence the level of free IGF and low levels of IGFBP-1 are associated with higher levels of IGF-1 (136).

High GI diets have also been associated with high density lipoprotein cholesterol (HDL-C), triglycerides, adiponectin and markers of oxidative stress. All of these effects may also be associated with breast cancer risk and suggest mechanisms that may link GI and breast cancer risk.

**Figure 2.7** Biologic plausibility

### 2.4 Conclusion

Research to date does not support a strong association between GI and breast cancer risk but limitations of study designs may influence the interpretation of the results. The investigation of GI and breast cancer risk has been confined to observational studies that used FFQs to collect dietary data. These studies have shown inconsistent results but a meta-analysis suggests that when consideration is given to the quality of the nutrient data, GI may be associated with breast cancer risk.

We are unable to draw any conclusions about the role of GI on breast cancer phenotype as only four studies stratified their results by ER/PR status and none considered nodal status or tumour size. The international variations in incidence of ER positive tumours, and the
effect of insulin on estrogen levels and cell proliferation, suggest that the effect of GI on phenotype may be important. Therefore, it would be worthwhile to further investigate an association between GI and breast cancer risk and phenotype with improved measurement of GI.

We propose to investigate the association of GI and breast risk and phenotype using the unique resource of a long term dietary intervention trial which will provide multiple food records over time. We have seen significant associations between total CHO, starch and total sugars in this population. Specifically, starch was inversely associated with ER positive breast cancer, and total CHO and total sugars were positively associated with risk of ER negative breast cancer. Dietary effects are broader than just the macronutrients and it will be important to look more specifically at different types of CHO foods and GI to understand the impact of carbohydrates on breast cancer risk. Food records are superior to FFQ in describing usual intake and therefore they will provide a better opportunity to detect a relationship between GI and breast cancer, if one exists. This work will also identify the specific types of CHO foods that influence breast cancer and allow the translation of this information into practical dietary advice.

2.5 Objectives

1. To calculate diet GI from multiple food records for subjects in a case control study (220 cases and 440 controls) nested within a trial cohort.

2. To determine if diet GI is associated with breast cancer risk overall and stratified by hormone receptor status.

3. To examine whether the observed associations of breast cancer risk and starch and total sugars are independent of the effects of diet GI.

4. To determine if diet GI is associated with breast cancer phenotype (hormone receptor status, tumor size and nodal status) in a case-only analysis.
Chapter 3 – Methods

The GI and Breast Cancer Risk and Phenotype Study used a nested case control design from the cohort of women who participated in the Diet and Breast Cancer Prevention Trial. Therefore, I will first provide an overview of the trial and then describe the methods specific to the GI study.

3.1 The Diet and Breast Cancer Prevention Trial

The Diet and Breast Cancer Prevention Trial (DBCPT) was the source of subjects for the GI and Breast Cancer Risk and Phenotype study. The trial began in 1988 and was a multi-centre, randomized, long term, intervention trial to determine if a low fat, high carbohydrate diet would reduce the incidence of breast cancer in healthy women at increased risk for the disease because they had extensive mammographic density. Study participants met regularly with a dietitian for 7 to 17 years and for an average of 10 years. They were asked to provide 3-day food records at baseline and every follow up visit and data on risk factors, anthropometric measures and a blood sample was collected at baseline and at every annual visit (see figure 3.1).

Figure 3.1 Design for the Diet and Breast Cancer Prevention Trial
Identification of Subjects

Potential participants were identified in mammography units on the basis of their mammogram, age and residence. Radiologists identified women with mammograms showing greater than 50% of the breast area occupied by dense breast tissue, and if the women were between 30 and 65 years of age and lived within commuting distance of the study centre they were asked to consider participating in the trial. However, to be eligible for the trial, women also had to meet seven additional eligibility criteria: 1. Not pregnant or breast feeding, or planning to be so within the first two years of randomization; 2. No history of cancer, except basal cell carcinoma of the skin, or any type of in situ cancer, except breast; 3. Serum cholesterol reported to be less than 6.2 mmol/L; 4. Not following a medically prescribed diet; 5. Body mass index between 19 and 27; 6. No history of breast reduction or augmentation; and 7. Willing to have a mammogram at least once every two years.

Screening

Women who were interested and eligible attended two screening visits. At the first screening visit a dietitian confirmed eligibility, collected baseline data, explained study procedures and taught how to keep food records. The dietitian completed a standardized health questionnaire to record educational level, occupation, medical history, reproductive history, use of medications, hormones and vitamins, as well as height (cm), weight (kg) and skin fold thickness (mm) at the triceps, subscapular and suprailiac. Potential participants were then asked to keep 3-day food records before attending the second screening visit which took place about two weeks later. At the second screening visit a dietitian reviewed the 3-day food record for clarity and completeness and took a comprehensive diet history to learn more about their food preferences, eating patterns, restaurant meals, special dietary needs and responsibility for shopping and cooking. If a woman was not following a low fat diet and demonstrated that she could keep appointments, provided adequate food records, prepared the majority of meals, ate in restaurants less than three times a week and was willing to accept randomization to either group, she was asked to provide consent.
Randomization

Subjects were randomized by telephone contact to the Biostatistics Department at the Ontario Cancer Institute and were informed of their group assignment at the baseline visit. By December 1998, we recruited 4,690 women from Toronto, Hamilton, Kitchener, London, Sarnia, Surrey, Vancouver and Windsor.

Dietary Intervention

The goal of the intervention was to achieve and sustain a low fat, high carbohydrate diet that would maintain body weight. The diet was based on an isocaloric replacement of fat with carbohydrate with a target energy distribution of 15% fat, 65% carbohydrate and 20% protein. In practical terms, the intervention restricted added and hidden fat, limited portions of lean animal protein foods and emphasized grains, vegetables and fruit.

At the initial instruction at baseline each subject received a handbook, called “The Fat Factor”, which included an individualized meal plan and sample menu to match their energy needs, food preferences and eating habits as well as restaurant guides, shopping guides and over 100 recipes. In addition, throughout the tenure of the trial we developed recipes on a monthly basis to further support the maintenance of a low fat, high carbohydrate diet.

Control Group

Subjects assigned to the control group were given a copy of Canada’s Food Guide for Healthy Eating at the baseline visit but were not asked to change their diet.

Follow up Schedule

All subjects met with a dietitian at each clinic visit. Subjects in the intervention group were seen monthly in the first year, quarterly in the second year, and every 6 months thereafter. Subjects in the control group were seen every 3 months in the first year, twice in the second year and then annually. All subjects were asked to bring 3 day food records to every follow up visit and they were reviewed by the dietitian for clarity and completeness. In addition, at every annual follow up visit, the dietitian completed a standardized health questionnaire like the form completed during their first screening visit, collected anthropometric measures (height, weight, waist and hip circumference, measured skin fold
thickness (triceps, subscapular and suprailiac) and asked the subjects to provide a blood sample.

**Nutrient Data**

Dietary intake was assessed using 3-day food records. The days were randomly selected by the dietitian to include two weekdays (working days) and one weekend day (non-working day). Participants were asked to describe foods and method of preparation in detail, including the type of fat used, if any. They were also asked to provide recipes for mixed dishes, brand names of crackers, cookies, cereal, candy and commercial entrees, the names of the specific cuts of meat, % MF of dairy products and the names of restaurants for meals eaten out. Measurements were recorded as a weight, volume, or cubic measurement. All study subjects were given a scale for weighing food for their food records. The volume was to be measured using standard measuring spoons and cups, and a ruler was to be used to determine the cubic dimensions of foods.

Food records were blinded for subject identity, study group and analyzed for nutrient content by trained dietitians using Nutrition Data System (NDS) version 2 from the Nutrition Coordinating Center, University of Minnesota. NDS contained some Canadian foods but the data entry process was modified to accommodate more Canadian foods and new commercial products. The nutrient data in NDS includes all macronutrients and many types of carbohydrate including: total dietary fibre, insoluble fibre, soluble fibre, starch, fructose, glucose, galactose, sucrose and lactose.

**Serum Measures**

Non-fasting blood samples were collected in the DBCPT and plasma/serum separated. Serum aliquots were frozen and stored -70° C.

**Study end points**

The DBCPT ended on December 31, 2005. At that time the minimum follow-up was 7 years and average follow-up was 10 years. In the 6 months before the end of the trial we contacted, or saw in person, 4,610 (98.3%) of 4,690 randomized subjects. No record of the remaining 80 subjects (1.7%) was found in the cancer registries of Ontario or British Columbia.
The primary end point was invasive breast cancer, and there were 220 cases in total, all were confirmed by pathology report. The pathology reports described estrogen and progesterone receptor status, tumor size and nodal status.

3.2 Glycemic Index and Breast Cancer Risk and Phenotype Study Design

This study used a nested case control design from the cohort of women who participated in the DBCPT (See Figure 3.2). The cohort was comprised of women from both the intervention and comparison groups of the trial. This design provides the advantages of a cohort study as the subjects come from the same population and the data on risk factors and dietary intake is collected prospectively before the diagnosis of breast cancer. The GI and Breast Cancer Risk and Phenotype study used data from the 220 cases of invasive breast cancer and 440 controls selected from the DBCPT. Each case was matched to two controls according to age at entry (within one year), trial centre, and time in the trial (date of recruitment within one year and active in the trial until the endpoint date of the case or longer).

![Figure 3.2](image)

**Figure 3.2** Design for the GI and Breast Cancer Risk and Phenotype Study

3.3 Data Collection

We used demographic, risk factor, anthropometric and dietary data as well as blood samples that were collected at baseline and annual visits from the 660 subjects in the case control study nested in the DBCPT. For this project we used food records provided at baseline and a minimum of three post-randomization visits. For cases diagnosed at or before
their Year 3 visit, we included all food records available from visits at month 4, month 8, year 1, month 18, year 2 and year 3. For cases diagnosed after three years of follow up, we used all available food records from annual visits. We selected the corresponding food records from their matched controls. There was an average of 20.5 (SD 10.4) food record days per subject. The median and mode was 18 food record days per subject. The number of food records ranged from 3 to 58 days, and the inter-quartile range was 14 to 27 food record days. The wide range in number of food record days was partly due to varying lengths of time in the trial until a diagnosis of breast cancer. Length of time in the trial explained 67% of variation in number of food record days for cases, and 33% of the variation for controls.

3.4. Development of the GI Database

3.4.1 Overview of GI Database
The versions of NDS we used in the DBCPT did not include GI values in their nutrient database. Therefore we developed a database of GI values that could be linked to the NDS codes and descriptors, rather than assigning GI values manually to each food item on thousands of food records. This linkage would add a GI value to the other nutrient data for each food item on each food record day. See Figure 3.3 for a flowchart of the method.

NDS had three descriptors for every NDS code: the general descriptor which related to the NDS code, the descriptor that represented the selections made at data entry for single foods, and the descriptor that represented the selections made at data entry for mixed foods or recipes. Since each NDS code was linked to more than one food item, it was very important to consider all three descriptors when assigning GI values.

Before we could assign GI values to the NDS codes and descriptors the data manager prepared an electronic file that contained nutrient data and descriptors for every food item on a food record. These files were exported to Microsoft Access and transformed into a food record database so we could refer to this information in order to determine the GI values. She prepared electronic files for batches of food records, starting with 50 records and gradually increasing the batch sizes to include 1,000 food records.

As we worked through the food items, it became apparent that in most cases the food items that were linked to the same NDS code had the same GI value. But we found 25 NDS codes that had multiple foods linked to them and the different foods required unique GI
values. We called these NDS codes “shared food codes”. We also considered the NDS code for corn syrup a shared food code because corn syrup was used in large amounts in a variety of foods including spaghetti sauce and soy milk, likely to raise their CHO content. If left unchanged, the GI associated with corn syrup (GI = 100) would raise the GI value of these foods substantially.

The database also served to document the decisions related to the assignment of each GI value. In addition to the GI value, for each NDS code we recorded: food group, type of match, source of the GI value, shared food code description, if needed, comments regarding rationale, date added to the database and date GI value was checked. See section 3.4.3 for a more detailed description of some of the specific fields in the GI database.

3.4.2 Source of GI Values

We predominantly used GI values published in scientific journals as the basis for assigning GI values to the foods in NDS. The primary source of GI values was the International Tables of Glycemic Index and Glycemic Load Values: 2008 compiled by FS Atkinson, K Foster-Powell and JC Brand-Miller (137). There were two tables of GI values: one with subjects with normal glucose tolerance, and one with subjects with impaired glucose tolerance or diabetes. The tables provided either references for their studies in the scientific literature or stated that they were unpublished observations of the Sydney University’s Glycemic Index Research Service. We checked the references for foods of interest and only considered studies that tested foods of North American formulation and used analytical methods consistent with the protocol recommended by the FAO/WHO expert consultation on carbohydrates in human nutrition (138). The Sydney University Glycemic Index Research Service in Australia followed the standard procedures for their measurements of GI values and we used their values if the foods tested were similar in carbohydrate composition to Canadian products. We also used GI values from two websites: President’s Choice Blue Menu website (www.pc.ca/bluemenuinfo) because it listed GI values for many Canadian foods that were tested by Glycemic Index Laboratories, Toronto, and the on-line database for The Sydney University Glycemic Index Research Service (www.glycemicindex.com) for any new values reported after the publication of the international tables. We used GI values based on the glucose standard.
Load NDS data files for food records into Food Record Database (food codes, descriptors and nutrient data)

Assign GI values to NDS Codes and descriptors in GI Database

Foods with ≤ 0.5 g ACHO/ serving

Assign GI = 0

Foods with > 0.5 g ACHO/ serving

Look for Direct Match in International Tables or on-line sources

Yes

Assign tested GI value

No

Look for similar food in tables or on-line sources

Yes

Assign tested GI value

No

Estimate GI value

Calculate GI using recipe for food items that are a mixture of foods

Assign GI value of dominant CHO in food

Link GI database with Food Record Data

Calculate diet GI

Figure 3.3 Flow chart of steps for calculating diet GI
3.4.3 Description of GI Database

a) **Food Group**

We assigned each NDS food code to a food group in order to facilitate a comparison of the GI values among similar types of foods and to allow for an examination of the relationship between the GI and other aspects of the diet.

b) **Type of Match**

We described the degree of match to a published value as:

i) “Direct” match when we matched to a comparable food with a published GI value.

ii) “Similar” match when we matched to a food that was very similar to a tested food. For example, we used the GI of white bread for white roll, tomato juice for tomato sauce, orange for mandarin orange.

iii) “Estimate” when we could not match to published values. The GI value was assigned based on the GI value of a comparable food, with similar type of carbohydrate and preparation method. For example, for ladyfinger cookies, which are like sponge cake, we used the mean GI value of 3 Canadian studies of sponge cake, angel cake and flan. In addition some foods with < 25 g of ACHO per typical serving were not able to be tested but would be expected to have some glycemic effect. In this situation, their GI value was based on the type of carbohydrate they contained (Jenkins 1981, Wolever 1985, 1994 from Schultz 2005). For example, organ meats and shell fish were assigned GI of glucose because they contain glycogen.

iv) “Recipe” for mixed foods. The GI value was calculated using the same methodology as the GI of a mixed meal. We assigned GI values to the ingredients in a mixed dish and then summed the product of the GI value for the ingredient and the proportion of ACHO of each ingredient in the recipe.

v) “Zero GI” when foods provided < 0.5 g ACHO per serving.
The bar graph below shows the distribution of NDS codes among the various types of matches.

**Figure 3.4** Distribution of NDS codes among type of matches with tested values

### c) Source

We documented the reference for the GI value or whether it was determined by the dietitians. The International Tables 2008 with GI values measured in healthy subjects was the source for 83% of NDS codes with a direct or similar match to a tested food.

### d) Shared Food Codes

We used two fields in the GI data base for shared food codes. One field was called “Food item for NDS Code” and contained a numeric value to indicate how many times the NDS code was shared. The second field was called “New Food Description” and contained the specific descriptor that required a different GI value.

### 3.4.4 Assignment of GI Values

The flowchart in figure 3.3 summarizes the method used for assigning GI values but a more detailed account of each step is described below:

**Step 1.** For foods with < 0.5 g ACHO per serving, we imputed GI value of zero. The following foods had GI value of zero: 1. sugar free soft drinks, 2. tea, coffee and alcohol, 3. butter, margarine, oil, shortening, mayonnaise and animal fats, 4. meat and poultry
except organ meats and processed meats, 5. fish except shellfish, 6. soy protein isolate, 7. yeast, gelatin powder, Tabasco, Worcestershire sauce, whole flax seeds and dried spices, and 8. mushrooms, watercress, nori and olives

**Step 2.** For foods with > 0.5 g ACHO per serving, we looked for a direct match in the International Tables or on-line sources. If more than one study was considered eligible, we calculated the simple mean of the GI values. If the food was not tested in subjects with normal glucose tolerance, we consulted the International Table of Glycemic Index and Glycemic Load Values: 2008 of studies using subjects with impaired glucose tolerance or type 2 diabetes.

**Step 3.** If the exact food could not be found, we checked for a closely related food in the GI Tables and assign the GI value for the similar food.

**Step 4.** If the food is not similar to tested product, we consulted the nutrition composition tables to determine the type(s) of carbohydrate and estimate GI value based on the closest match to a tested food with a similar composition and preparation method. We referred to McCance and Widdowson’s The Composition of Foods (139), USDA online database, Canadian Nutrient File online database and food labels to provide information on carbohydrate content of foods.

**Step 5.** If the food item is a mixture of foods, then we calculated the GI as a recipe. We identified the ingredients in the food; determined GI for each ingredient and calculated the overall GI according to a weighted average of the GI values for each ingredient, based on the proportion of the total CHO contributed by each ingredient (140).

**Step 6.** For low carbohydrate foods (< 25 g ACHO per serving) that were not tested, we imputed GI value corresponding to the dominant carbohydrate (monosaccharide, disaccharide, starch or high fructose corn syrup) in the product.
3.4.5 Considerations in the Assignment of GI Values

In many cases, multinational corporations, like Kellogg’s, used different formulations for their products in different countries. Therefore, when assigning GI values to Canadian products comparisons to tested foods from Australia or Europe were based on ingredient and nutrient information from labels or company websites rather than product name alone.

Specific considerations for assigning GI values to different types of foods are summarized below:

**Alcoholic Beverages – Cocktails**

We used the GI of grapes for wine. We used the GI of sugar for liqueurs sweetened with sugar and calculated the GI of cocktails using a recipe based on their usual ingredients.

**Baked Goods**

We calculated the GI of cakes, pies, muffins and non-frozen desserts by recipe, using the GI of white bread for flour.

**Candy and Chocolate**

Studies of candy and chocolate bars with nuts tend to have a lower GI than expected by calculation using a recipe. Nuts are very low GI foods and seem to exert a significant GI lowering effect on these foods as evidenced by studies of Munch Bar (GI=27), Nutella (GI = 25, 30, 33) and M&M Peanuts (GI = 33). Therefore, we used the GI of Munch Bar (GI= 27) for candy that was mostly nuts. We matched other chocolate bars to one of the other three chocolate bars that have been studied: 1. Milk Chocolate (GI = 43) 2. Snickers Bar (GI=51) and 3. Mars Bar (GI=65).

The GI of clear candies like gum drops and life savers was based on the mean of five Australian studies of clear candy - jelly beans (76, 80), licorice (78), lifesavers (70) and skittles (70) since there were no studies conducted in North America.

**Cereal**
There were a small number of studies of Canadian breakfast cereal. We estimated GI values for untested cereals based on the closest match to a tested product giving consideration to the type of grain (wheat, oat, rice, corn), form (puffed, nugget, flake) and type of sweetener (sugar, honey or HFCS).

**Cookies**

Most cookies have similar amounts of flour, sugar and fat and very little liquid and are prepared in a similar way. Consequently, we expect the GI values of most cookies to be similar. According to Englyst, biscuits tend to have lower GI values than crackers and cereals because they have more slowly absorbed glucose due to lower moisture content and less manipulation during processing (141).

There were few studies of cookies and therefore we used the mean of tested plain cookies (digestive, arrowroot, social tea and oatmeal) for all cookies except chocolate chip, chocolate coated, fruit filled and peanut butter cookies. Fruit filled cookies were tested and matched directly to NDS codes. The GI values of chocolate chip and chocolate coated cookies were calculated by recipe of plain cookie and chocolate.

** Crackers **

Most crackers are made of wheat flour, water and fat with some seasoning. The production method for crackers is fairly consistent therefore the degree of gelatinization of the starch in flour is similar. Therefore we used the mean of tested crackers for all wheat based crackers. Melba toast is dried bread so we used the GI of white bread. We used tested values for rye crackers and graham crackers.

**Frozen Desserts**

Our decisions about the GI of frozen desserts were based on information about the production of ice cream, frozen yogurt and sherbet from the Dairy Science and Technology website of the Food Science department of the University of Guelph (www.foodsci.uoguelph.ca/dairysci) and from studies of North American products. In Canada, regular ice cream and frozen yogurt are sweetened with sugar and glucose and premium ice cream is sweetened with sugar. Soft ice cream is similar to regular ice cream
but transferred to bags before hardening. The measured GI of frozen desserts tended to be lower than expected by calculation of a recipe perhaps because of the effect of cold extrusion (128). Therefore, when we used a recipe to determine the GI of a frozen dessert we reduced the GI by 20% if cold extrusion was used in its production.

Fruit

If a GI value was not available for a fresh, dried or canned fruit, or juice, we consulted food composition tables to determine their carbohydrate composition and imputed the GI value for the tested fruit or juice with a similar profile. We applied the same GI for all berries which was the mean of two studies: Strawberries (GI 40 - Australia) and Blueberries (GI 53 - Canada). All berries have a similar carbohydrate content which is an almost equal amount of glucose and fructose and very little sucrose.

We determined the GI for untested, sweetened fruits by recipe calculation. This procedure was applied to baked apples, apple chips, spiced apples, frozen sweetened berries, fruit nectars, cranberry sauce and fruit canned in heavy syrup.

Although we expected the GI of juice to be higher than the GI of its fruit, it is difficult to predict the GI of juice from the GI of the fruit it is made from. The studies of apples and apple juice, suggest a 5% increase, and studies of oranges and orange juice suggest a 25% increase. These results made it impossible to estimate the GI of juice if it was not tested. In these cases, rather than guess, we used the tested value of the fruit.

High Fructose Corn Syrup

There are two types of high fructose corn syrup (HFCS) used in commercial production of beverages and food. HFCS 55%, which is 55% fructose and 45% glucose, is used to sweeten soft drinks. Therefore, we used the tested value for cola (GI=63) for the GI of HFCS 55%. The GI of cola is, as expected, similar to the GI of sucrose (GI=61) as they contain similar amounts of glucose and fructose. HFCS 42%, which is 42% fructose and 58% glucose, is ubiquitous in the marketplace as a sweetener for a wide variety of other foods. The tested value of HFCS 42% (GI=81) (142) is considerably higher than the calculated value (GI=66) based on 42/100 g fructose with GI = 20 and 58/100 g glucose with GI = 100. Therefore, we used the mean of the tested and calculated values for HFCS 42% (GI =74).
**Milk, Cheese and Cream**

Changes in fat content in milk do not affect the GI. We used the mean of Canadian and US tests of all types of milk for skim, 1/2%, 1%, 2% and whole milk. Since the only carbohydrate in cheese and cream is lactose, we applied the GI of milk to hard, cream and cottage cheese, non-flavoured whipping cream and sour cream. The GI of sweetened cream products was calculated by recipe of milk and type of sweetener, sugar or HFCS.

**Nuts and Seeds**

Cashews were the only nuts tested for GI value. According to food composition tables, all nuts have similar amounts of protein and fat, and the carbohydrate is consistent and almost exclusively sucrose, we imputed the mean tested GI value for cashews for all nuts. Sunflower and sesame seeds were never tested. They are similar to nuts in their nutrient composition, and we therefore used the GI value for cashews for seeds.

**Potatoes**

There have been many studies of baked and boiled potatoes which produced a wide range of GI values (56 to 98). The variation in GI values did not appear to be related to cooking method or variety of potato. Therefore, we used the mean of all studies of baked or boiled potatoes that followed the standardized procedures for measuring GI for boiled, baked, microwaved and mashed potatoes (mean GI = 77).

Resistant starch is present in raw potatoes and is mostly converted to digestible starch on cooking. The starch in potatoes will retrograde when chilled, resulting in a lower GI value for potato salad. Similarly, as there would also be some retrogradation of the starch in frozen potato products, as they are chilled and then reheated, we used the GI of frozen French fries for other frozen potato products like hash browns and potato croquettes. Instant mashed potatoes are very highly processed to create a dried, flaked product that is easily hydrated and consequently have the highest GI value of all potatoes tested.
Processed Meats

Processed meats often contain small amounts of sugar, corn syrup and/or starch. Therefore, we assigned a GI value based on the commonly used filler or sweetener in the particular style of processed meat.

Rice

There have been many studies of white, long grain rice conducted in North America, Europe, Asia and Australia and they produced a wide range of GI values (38 to 93). This may be due to the wide variety of rice available around the world with varying amounts of amylase and amylopectin. Since the type of long grain rice was not usually specified in the GI tables, and we do not know the specific type of white rice our subjects ate, we used the value from the Canadian study (GI = 69) as it would likely be most representative of the type of rice our subjects might have eaten.

Vegetables

A GI value was not available for many vegetables because they have small amounts of ACHO and subjects would have to consume very large amounts in order to test the glycemic effect of at least 25 g ACHO. Therefore, we consulted food composition tables to determine the carbohydrate composition of each untested vegetable and imputed the GI value for a corresponding tested vegetable with a similar profile of carbohydrates. The GI was measured in beets, broad beans, carrots, corn, peas, parsnips, pumpkin, sweet potatoes, tomato juice and vegetable juice. We used the GI of tomato juice for most low carbohydrate vegetables and fresh herbs because it was the only vegetable tested that shared the same distribution of carbohydrate, namely mostly glucose and fructose, with little sucrose and no starch.

Yogurt

Like milk, the GI value of yogurt is not affected by fat. However, the formulation of yogurt in the UK and Australia is different than in Canada and this variation did influence the GI value. We based our determination on Canadian studies and decided on two GI values for yogurt. One GI value was assigned to plain, unsweetened yogurt, or yogurt sweetened with a
sugar substitute (GI = 20). Another GI value was used for all flavours of sweetened yogurt, including fruit, vanilla or coffee flavours, since sugar was the predominant sweetener in all of these products (GI = 40).

3.5 Quality Control

Experience from the European Prospective Investigation into Cancer and Nutrition Trial indicated that there can be considerable inter-rater variation in the assignment of GI values (143). Therefore, Ms. Gougeon and I collaborated in the assignment of GI values to the NDS codes and descriptors. We reached a consensus on every GI value, regardless of the type of match to a tested value: direct, similar or estimated. We documented our decision in our GI database and maintained a paper record of GI values according to NDS code. As we worked through the food record databases to assign GI values, we compared values assigned to new NDS codes, to the values already assigned to similar foods. This step ensured we were consistent or could justify any deviations from prior decisions.

After the GI database was completed, the GI values were checked by food group. The GI database was sorted by food group and in ascending order of GI values. Any outliers or inconsistencies were verified by checking our written records.

The data manager checked the GI values for food items with shared food codes electronically to identify any unexpected values and I verified every irregularity.

3.6 Calculating Diet GI

The final database of NDS codes and GI values were linked electronically with food record data. The diet GI for each food record was calculated by the formula described by Wolever et al (144). The overall diet GI was defined as the weighted mean of the GI of individual CHO containing foods and is calculated as: $\frac{\sum (\text{available CHO in g for each food} \times \text{GI for each food})}{\text{total available CHO in g/day}}$. We calculated the diet GI at baseline by averaging the diet GI from the baseline food records, and then calculated the diet GI post-randomization by averaging the diet GI from all post-randomization records.
3.7 Pilot Study to compare automated vs manual system

Prior to assigning GI values to the food records in the case-control study, we conducted a pilot study of an independent sample of 50 subjects with one food record day per subject. We compared GI values determined using the automated method of assigning GI values from the electronic food record (described in section 3.4) with the manual method of assigning GI values from the actual food record. Ms. Gougeon and I collaborated in the assignment of GI values for both methods to avoid inter-rater variation.

The data manager used information from NDS, specifically from the data entry of each item on the food records and the results of the nutrient analysis, to produce a food record database. It contained all the food items on the 50 food records and also had two fields for GI values to correspond to the automated and manual method. For the automated method, GI values were determined by an electronic linkage to the GI database. For the manual method, the dietitian used the handwritten food record which provided additional information about foods, such as brand names and recipes for mixed dishes that might influence the determination of the GI values. If the information on the food record suggested a different GI value then it was recorded in a separate field in the food record database.

We calculated the diet GI derived from the two methods. The Diet GI calculations from the two methods were very highly correlated ($r = 0.98$) which confirms that the automated system was able to capture the information needed determining the diet GI using an automated system is feasible (see Figure 3.5 for the correlation).

![Figure 3.5 Correlation of manual and automated method for determining diet GI values](image)

$r = 0.98$
3.8 Statistical Methods

Descriptive statistics included a comparison of subject characteristics and nutrient intakes by case control status. I used Student’s t test for continuous variables with approximately normal distributions, Wilcoxon rank sum test for continuous variables whose distributions were skewed, and chi-square test for categorical variables. Nutrient intakes were analyzed at baseline and post-randomization and the mean difference was tested using Student’s t test when changes in nutrient intake were normally distributed and the Wilcoxon rank sum test when the distributions were skewed.

A comparison of diet GI by subject characteristics were conducted using Student’s t test for continuous variables with approximately normal distributions, Wilcoxon rank sum test for continuous variables whose distributions were skewed, and chi-square test for categorical variables. Correlation analysis was used to assess the relationships between diet GI and nutrients.

The association of diet GI with breast cancer risk was examined using conditional logistic regression with case/control status as the dependent variable and diet GI as the independent variable, before and after controlling for potential confounders such as breast cancer risk factors, study group, weight, energy intake and other selected nutrients. Analysis was carried out for all invasive breast cancers and stratified by hormone receptor status.

The association of diet GI with receptor and nodal status was examined using unconditional logistic regression in breast cancer cases only with status (positive/negative) as the dependent variable, and diet GI or glycemic index as the independent variables, before and after controlling for potential confounders (see above). The association of GI with tumour size was assessed in a similar manner using multiple linear regression. We can detect a difference of at least 3.75 GI units between tumours positive and negative for ER (OR of at least 1.63 between upper and lower quartiles), and a difference of at least 3.1 units (OR of 1.50) between tumours with positive and negative nodal status.

We have estimated the minimal detectable difference in mean GI between cases (n=220) and controls (n=440) based on a t-test (2 sided, α=0.05; 80% power). Assuming mean GI (55.1), SD (5.3) and interquartile range (5.0) as observed in our pilot data, we can detect a difference of at least 1.72 units of GI (corresponding to an OR of 1.25 between top and bottom quartiles). In the subsets of ER positive and negative tumours, ORs of 1.3 and 1.7
respectively can be detected.

3.9 Candidate’s Contribution to the Project

I have been the nutrition coordinator of the DBCPT since 1992 and during the operation of the trial I was responsible for and participated in the collection of demographic, anthropometric, medical and nutrient data used in this study. From 1992 to 2007 I entered thousands of food records into NDS for the DBCPT and specifically for the nested case-control study which was used in this project. Dr. Lisa Martin developed the overall study design for this project and I developed the procedures and method for assigning GI values to the NDS codes and calculating the diet GI values. The data manager and I designed the GI and food record databases. I collected the published and on-line sources of GI values and hired and trained a dietitian with experience in nutrient data entry and extensive knowledge of food composition to collaborate in assigning GI values to the NDS codes. We worked together to reach a consensus on every GI value in our database. I conducted the quality control reviews of the GI database. The data manager prepared the food record data base, linked the GI database with the NDS nutrient data and electronically calculated the diet GI values for each food record day. I did the data analysis for the descriptive statistics, and the comparisons of group means, correlations and linear regressions that describe the characteristics and dietary intake of the study population, the associations of diet GI with case control status, subject characteristics and nutrient intake, the correlations of diet GI with nutrients and the linear regression model to predict it.
Chapter 4: Results

The results are presented in four sections: (1) descriptive statistics about subject characteristics and nutrient intake; (2) the association of case control status, subject characteristics and nutrients with diet GI using univariate analysis; (3) the relationship of diet GI with breast cancer risk using conditional logistic regression and (4) the relationship of diet GI with tumor phenotype in cases only using logistic regression.

4.1 Descriptive Statistics

4.1.1 Characteristics of Subjects

The distribution of risk factors among the 220 cases and 440 controls in our nested case-control study at the time of entry to the DBCPT is shown in Table 4.1. As expected, because subjects were matched by age, there was no difference in age between cases and controls (mean of 48.5). Seventy-two percent of cases and 68 % of controls were premenopausal at entry to the trial ($\chi^2 = 1.15, p = .28$). The mean BMI was in the healthy range for cases (mean 23.7, SD 2.7) and controls (mean 23.4, SD 2.4). Weight, height, age at menarche, parity, age at first live birth, number of children for parous women, menopausal status and hormone ever use were not significantly different between cases and controls. Within cases and controls there was a similar distribution of subjects from the intervention and control groups of the DBCPT (p = 0.30). Cases were more likely to have a family history of breast cancer, 23 % of cases had a first degree relative with breast cancer compared to 17% of controls (p =0 .04). Most subjects were married (76%), completed college or university (54%) and never smoked (50%). There were no significant differences in marital status, education level or smoking status between cases and controls.
### Table 4.1 Selected baseline characteristics of subjects in the nested case-control study

<table>
<thead>
<tr>
<th></th>
<th>Cases a (n = 220)</th>
<th>Controls b (n = 440)</th>
<th>P value a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD) or %</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>48.5 (6.3)</td>
<td>48.5 (6.2)</td>
<td>1.00</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>63.4 (8.9)</td>
<td>62.2 (7.3)</td>
<td>0.08</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>163.5 (6.4)</td>
<td>163.1 (5.8)</td>
<td>0.42</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.7 (2.7)</td>
<td>23.4 (2.4)</td>
<td>0.15</td>
</tr>
<tr>
<td>Age at Menarche (years)</td>
<td>12.8 (1.3)</td>
<td>13.0 (1.5)</td>
<td>0.17</td>
</tr>
<tr>
<td>Parity (% parous)</td>
<td>65.9</td>
<td>70.5</td>
<td>0.23</td>
</tr>
<tr>
<td>Age at First Live Birth (years)</td>
<td>26.1 (5.2)</td>
<td>25.7 (4.9)</td>
<td>0.41</td>
</tr>
<tr>
<td>Number of Children for Parous Women</td>
<td>2.0 (1)</td>
<td>2.0 (1)</td>
<td>0.40</td>
</tr>
<tr>
<td>Menopausal Status (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premenopausal</td>
<td>71.8</td>
<td>67.7</td>
<td>0.28</td>
</tr>
<tr>
<td>Postmenopausal</td>
<td>28.2</td>
<td>32.3</td>
<td></td>
</tr>
<tr>
<td>Hormone Ever Use (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>31.4</td>
<td>28.2</td>
<td>0.40</td>
</tr>
<tr>
<td>No</td>
<td>68.6</td>
<td>71.8</td>
<td></td>
</tr>
<tr>
<td>First Degree Relative with Breast Cancer (% yes)</td>
<td>23.2</td>
<td>16.6</td>
<td>0.04</td>
</tr>
<tr>
<td>Study Group (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>53.6</td>
<td>49.3</td>
<td>0.30</td>
</tr>
<tr>
<td>Control</td>
<td>46.4</td>
<td>50.7</td>
<td></td>
</tr>
<tr>
<td>Marital Status (%)</td>
<td>Cases (n = 220)</td>
<td>Controls (n = 440)</td>
<td>P-value&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Never Married</td>
<td>12.7</td>
<td>13.9</td>
<td>.80</td>
</tr>
<tr>
<td>Divorced</td>
<td>7.7</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>Separated</td>
<td>2.3</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>75.9</td>
<td>76.4</td>
<td></td>
</tr>
<tr>
<td>Widow</td>
<td>1.4</td>
<td>2.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education Level (%)</th>
<th>Cases (n = 220)</th>
<th>Controls (n = 440)</th>
<th>P-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than High School</td>
<td>9.5</td>
<td>5.5</td>
<td>.19</td>
</tr>
<tr>
<td>High School or Technical School</td>
<td>21.8</td>
<td>26.1</td>
<td></td>
</tr>
<tr>
<td>College or University</td>
<td>54.1</td>
<td>53.2</td>
<td></td>
</tr>
<tr>
<td>Post-Graduate or Professional School</td>
<td>14.5</td>
<td>15.2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Smoking Status (%)</th>
<th>Cases (n = 220)</th>
<th>Controls (n = 440)</th>
<th>P-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never Smoked</td>
<td>52.3</td>
<td>50.2</td>
<td>.60</td>
</tr>
<tr>
<td>Past Smoker</td>
<td>37.7</td>
<td>41.4</td>
<td></td>
</tr>
<tr>
<td>Current Smoker</td>
<td>10.0</td>
<td>8.4</td>
<td></td>
</tr>
</tbody>
</table>

**Legend:**

<sup>a</sup> P value from 2 sample t-test for: age, weight, height, BMI, age at menarche, age at first live birth; Wilcoxon Mann-Whitney test for number of children born to parous women; Chi square test for categorical variables: parity, family history, menopausal, hormone ever use, study group, marital status, education level, smoking status and study site.

<sup>b</sup> Data missing for age at menarche: 1 case, 2 controls

<sup>c</sup> Data missing for age at first birth: 1 control

<sup>d</sup> Number of Children for Parous Women reported median and IQR
4.1.2 Nutrient Intake

The average energy and nutrient intakes for subjects in the nested case control study came from analysis of baseline and post-randomization (collected after baseline) food records from the DBCPT. As the subjects for this study came from a dietary intervention trial, changes in energy, weight and nutrient intakes occurred for some subjects and these are shown below. Since I examined diet GI in this context, I have conducted separate analyses for baseline and post-randomization.

4.1.2.1 Energy and Macronutrients

Table 4.2 shows energy, weight, and reported intake of macronutrients for cases and controls. Baseline energy, weight and intakes of macronutrients were similar in the cases and controls, except for small but statistically significant differences in total fat. Cases consumed more total fat (mean 58.6 g, SD 21.3) compared to controls (mean 55.2 g, SD 19.4) (p = 0.04). After randomization, there was no significant difference between cases and controls in macronutrient intakes except for total protein. Cases consumed more total protein (69.7 g, SD 13.4) than controls (66.8 g, SD 11.6) (p = 0.008) but there was no difference in percent energy from protein.

4.1.2.2 Types of Carbohydrate

Table 4.3 shows intake of types of carbohydrates for cases and controls: total fibre, available CHO, starch, sucrose, lactose, fructose and total sugars. Available CHO (ACHO) is calculated by total CHO minus total fibre and represents the amount of starch and total sugars. Total sugars are calculated by ACHO minus starch and represent the amount of all mono- and disaccharides. Baseline intakes of types of carbohydrate were similar between cases and controls except for a small but statistically significant difference in total sugars (p = 0.03) which were primarily due to a significant difference in sucrose. Cases consumed more sucrose (mean 37.9 g, SD 19.4) than controls (mean 34.4 g, SD 16.1), p = 0.02. There were no significant differences in consumption of types of carbohydrate between cases and controls at post-randomization. However, cases consumed more fructose (mean 23.3, SD 9.0) and total sugars (mean 114.1, SD 34.4) than controls (mean fructose 22.0, SD 8.4; mean total sugars 109.1, SD 32.9), the difference was not significant (p = 0.08).
Table 4.2  Energy, weight and macronutrient intake

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Case</th>
<th>Baseline</th>
<th>Post-randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Status</td>
<td>220 cases, 440 controls</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>P value b</td>
</tr>
<tr>
<td>Energy (kcal)</td>
<td>Case</td>
<td>1683</td>
<td>414.7</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>1620</td>
<td>375.6</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>Case</td>
<td>63.4</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>62.2</td>
<td>7.3</td>
</tr>
<tr>
<td>Total Fat (g)</td>
<td>Case</td>
<td>58.6</td>
<td>21.3</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>55.2</td>
<td>19.4</td>
</tr>
<tr>
<td>Protein (g)</td>
<td>Case</td>
<td>67.5</td>
<td>16.9</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>65.6</td>
<td>16.2</td>
</tr>
<tr>
<td>Total CHO (g)</td>
<td>Case</td>
<td>216.0</td>
<td>60.5</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>212.7</td>
<td>57.7</td>
</tr>
<tr>
<td>Alcohol (g) d</td>
<td>Case</td>
<td>4.0</td>
<td>12.8</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>3.1</td>
<td>10.4</td>
</tr>
<tr>
<td>% Fat</td>
<td>Case</td>
<td>30.8</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>30.0</td>
<td>6.9</td>
</tr>
<tr>
<td>% Protein</td>
<td>Case</td>
<td>16.5</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>16.6</td>
<td>3.3</td>
</tr>
<tr>
<td>% Total CHO</td>
<td>Case</td>
<td>51.7</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>53.0</td>
<td>8.2</td>
</tr>
</tbody>
</table>

Legend:

a  Data missing for pos-randomization nutrients: 12 cases, 10 controls
b  P value from 2 independent samples t-test for: baseline and post-randomization energy, weight, total fat, protein, total CHO, % fat, % protein, % total CHO; and baseline total fat. Wilcoxon Mann-Whitney test for: baseline and post-randomization alcohol
P value from paired samples t-test for: energy, weight, total fat, protein, total CHO, % fat, % protein, % total CHO. Wilcoxon Signed Ranks two related samples test for alcohol.
c  Data missing for post-randomization weight: 1 case, 8 controls
d  Median difference reported for alcohol
### Table 4.3 Intake of types of carbohydrate

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Case</th>
<th>Control</th>
<th>Baseline 220 cases, 440 controls</th>
<th>Post-Randomization 208 cases, 430 controls&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>P value&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Mean</td>
</tr>
<tr>
<td>Total Fibre (g)</td>
<td>Case</td>
<td>18.9</td>
<td>7.4</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>18.7</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>ACHO (g)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Case</td>
<td>197.1</td>
<td>56.1</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>194.0</td>
<td>53.4</td>
<td></td>
</tr>
<tr>
<td>Starch (g)</td>
<td>Case</td>
<td>91.8</td>
<td>30.7</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>95.1</td>
<td>30.9</td>
<td></td>
</tr>
<tr>
<td>Sucrose (g)</td>
<td>Case</td>
<td>37.9</td>
<td>19.4</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>34.4</td>
<td>16.1</td>
<td></td>
</tr>
<tr>
<td>Lactose (g)</td>
<td>Case</td>
<td>13.5</td>
<td>8.5</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>13.1</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Fructose (g)</td>
<td>Case</td>
<td>20.2</td>
<td>9.7</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>19.3</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td>Total Sugars (g)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Case</td>
<td>105.3</td>
<td>36.0</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>98.9</td>
<td>34.7</td>
<td></td>
</tr>
</tbody>
</table>

**Legend:**

<sup>a</sup> Data missing for post-randomization nutrients: 12 cases, 10 controls

<sup>b</sup> P value from 2 independent samples t-tests

<sup>c</sup> P value from paired samples t-tests

<sup>d</sup> ACHO = total carbohydrate minus total fibre

<sup>**</sup> Total Sugars = ACHO minus starch
4.2 Diet GI

The distribution of diet GI for cases and controls at baseline, post-randomization and for the overall mean of all food records are shown in the histograms in figure 4.1. The distributions appear symmetrical for both groups at all time points. The mean and median are virtually identical among cases at baseline (mean 57.6, SD 3.2, median 57.7), post-randomization (mean 57.4, SD 2.5, median 57.7) and with the overall mean (mean 57.6, SD 2.5, median 57.6). The mean and median diet GIs are also virtually identical among controls at all time points. The data appeared normally distributed and was not transformed for analysis.

The range of diet GI values as described by the IQR was 55-60 for cases and 56-60 for controls at baseline; and at post-randomization and with the overall mean the IQR was 56-59 for both cases and controls. Other studies of diet GI and breast cancer risk described the range of GI values in their population as the mean of the 1st and 5th quintiles. In our study, using the data from controls, the mean of the 1st and 5th quintiles was 54-60. The range of GI values was similar between cases and controls and remained unchanged over time.

The mean differences in diet GI between baseline and post-randomization in cases and controls were examined using a paired samples t-test. The mean difference in diet GI among 207 cases was 0.16 (SD 2.9) and this change was not significant (t = 0.78, p = 0.44). The mean difference in diet GI among 430 controls was 0.12 (SD 2.8) and this change was also not significant (t = 0.86, p = 0.39). Therefore, there were no significant changes in diet GI over time in either group.

Since the diet GI was consistent at all time points, the best estimate of usual diet GI would likely be the overall mean of all available food records because it is based on the most food record data. However, in view of the changes in the intakes of other nutrients over time, I did the analysis for the associations of diet GI and breast cancer risk at baseline, post-randomization and with the overall mean of all food records.
Figure 4.1 Histograms of diet GI among cases and controls
4.2.1 **Comparison of Diet GI by Case Control Status**

The box plots of diet GI for cases and controls, at baseline, post-randomization and for overall mean of all food records are shown in figure 4.2. The difference in diet GI between cases and controls was tested using an independent two sample t-test. There was no difference in baseline diet GI between 220 cases (mean 57.6, SD 3.2) and 440 controls (mean 57.5, SD 3.1), $t = 0.72, p = 0.47$. Similarly, there was no difference in post-randomization diet GI between 207 cases (mean 57.4, SD 2.5) and 430 controls (mean 57.3, SD 2.2), $t = 0.45, p = 0.65$; and there was no difference in mean of all food records between 220 cases (mean 57.6, SD 2.5) and 440 controls (mean 57.4, SD 2.2), $t = 1.06, p = 0.29$. When the potential outliers, as seen in the box plots outside the whiskers, were removed from the analysis there were no changes in the results.

Further analysis of case control differences in diet GI using conditional logistic regression which accounts for matching and adjusts for breast cancer risk factors is described in section 4.3.
Figure 4.2 Box plots of Diet GI at (A) baseline, (B) post-randomization and (C) mean of all food records by case control status.
4.2.2 **Associations of Diet GI by Subject Characteristics**

The associations of diet GI with subject characteristics in cases and controls, at baseline and post-randomization, are shown in Table 4.4. The top section shows correlations with continuous variables, and the bottom section shows associations with categorical variables.

Age was negatively associated with baseline diet GI in both cases ($r = -0.18$, $p = 0.007$) and controls ($r = -0.13$, $p = 0.005$) and with post-randomization diet GI in cases ($r = -0.22$, $p < .0001$). There was no association between diet GI and weight, height, BMI, age at menarche, age at first live birth or number of children for parous women, among cases or controls at either time point.

The associations between diet GI and parity, menopausal status, HRT use, family history and DBCPT study group were tested using independent samples t-tests. Premenopausal women had a higher diet GI than postmenopausal women. Among cases, premenopausal women had a significantly higher baseline diet GI than postmenopausal women (mean difference = 1.12, SE = 0.49, $p = 0.03$). Also, among controls at post-randomization, pre-menopausal women had a significantly higher diet GI than post-menopausal women (mean difference = 0.63, SE = 0.24, $p = 0.008$). This association with menopausal status is consistent with the effect of age on diet GI. There were no significant differences in diet GI between subjects who were: parous or non-parous, HRT users or non-users, from the intervention or control groups, or between subjects with or without a first degree relative with breast cancer, among cases or controls, at either time point.

The association of diet GI with marital status and smoking status was tested using one way analysis of variance. There were no significant differences in diet GI among cases or controls by marital status or smoking status at either time point (data not shown).

The relationship between education level and diet GI was examined by comparing the difference in mean diet GI at baseline and post-randomization among cases and controls between 4 groups of subjects who completed: (1) less than high school, (2) high school or technical school, (3) college or university and (4) post graduate school. A one-way analysis of variance revealed that there was a significant difference in baseline diet GI among controls ($F=3.46$, $p = 0.02$). The difference between specific groups was tested using the Bonferroni post hoc test but there was no statistically significant difference. There was a similar trend
among cases, but the difference in baseline diet GI between groups was not significant (F = 2.43, p = 0.07). There was no significant difference in post-randomization diet GI between education levels among cases (p = 0.24) or controls (p = 0.26) (data not shown).

The association of diet GI by characteristics of the subjects was also conducted using the overall mean diet GI and the results were similar to post-randomization diet GI (data not shown).
Table 4.4 Associations of diet GI with characteristics of subjects

<table>
<thead>
<tr>
<th></th>
<th>Case</th>
<th>Control</th>
<th></th>
<th></th>
<th>Case</th>
<th>Control</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Status</td>
<td>Correlation Coefficient</td>
<td>P value a</td>
<td>Correlation Coefficient</td>
<td>P value a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline</td>
<td>Post-randomization</td>
<td></td>
<td>Baseline</td>
<td>Post-randomization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>Case</td>
<td>-0.18</td>
<td><strong>0.007</strong></td>
<td></td>
<td>-0.08</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>-0.13</td>
<td><strong>0.005</strong></td>
<td></td>
<td>-0.22</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg) b</td>
<td>Case</td>
<td>0.07</td>
<td>0.28</td>
<td></td>
<td>0.13</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.05</td>
<td>0.26</td>
<td></td>
<td>0.05</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height (cm)</td>
<td>Case</td>
<td>-0.03</td>
<td>0.68</td>
<td></td>
<td>0.06</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.03</td>
<td>0.50</td>
<td></td>
<td>0.05</td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²) b</td>
<td>Case</td>
<td>0.10</td>
<td>0.12</td>
<td></td>
<td>0.11</td>
<td>0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.04</td>
<td>0.46</td>
<td></td>
<td>0.03</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at Menarche (years)</td>
<td>Case</td>
<td>0.03</td>
<td>0.62</td>
<td></td>
<td>0.04</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.06</td>
<td>0.20</td>
<td></td>
<td>0.08</td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at First Live Birth for parous women (years)</td>
<td>Case</td>
<td>-0.10</td>
<td>0.26</td>
<td></td>
<td>-0.02</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.03</td>
<td>0.66</td>
<td></td>
<td>-0.07</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Children for Parous Women</td>
<td>Case</td>
<td>-0.06</td>
<td>0.50</td>
<td></td>
<td>0.007</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>-0.05</td>
<td>0.40</td>
<td></td>
<td>-0.05</td>
<td>0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity (children vs no children)</td>
<td>Case</td>
<td>0.56 (.46)</td>
<td>0.22</td>
<td></td>
<td>0.15 (.36)</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.51 (.33)</td>
<td>0.12</td>
<td></td>
<td>0.21 (.24)</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family History (yes vs no)</td>
<td>Case</td>
<td>-0.11 (.52)</td>
<td>0.84</td>
<td></td>
<td>-0.53 (.42)</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.26 (.41)</td>
<td>0.53</td>
<td></td>
<td>0.27 (.32)</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Menopausal Status (pre- vs post-)</td>
<td>Case</td>
<td>1.12 (.49)</td>
<td><strong>0.03</strong></td>
<td></td>
<td>0.33 (.42)</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.44 (.31)</td>
<td>0.16</td>
<td></td>
<td>0.63 (.24)</td>
<td><strong>0.008</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRT ever use (yes vs no)</td>
<td>Case</td>
<td>-0.85 (.48)</td>
<td>0.08</td>
<td></td>
<td>-0.03 (.38)</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.26 (.33)</td>
<td>0.43</td>
<td></td>
<td>0.04 (.26)</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Group (intervention vs control)</td>
<td>Case</td>
<td>-0.17 (.43)</td>
<td>0.69</td>
<td></td>
<td>0.44 (.35)</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>0.26 (.29)</td>
<td>0.38</td>
<td></td>
<td>-0.11 (.21)</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Legend:

a P value from Pearson correlations for: age, weight, height, BMI, age at menarche, age at first live birth, number of children for parous women; independent samples t-tests for: parity, family history, menopausal status, HRT ever use and study group.

b Correlations between post-randomization weight and BMI with post-randomization diet GI

c Mean difference calculated as:
Parity: mean GI of parous women – mean GI of non parous women
Family History: mean GI of women with family history – mean GI of women without family history
Menopausal Status: mean GI of premenopausal women – mean GI of postmenopausal women
HRT ever use: mean GI of HRT ever users – mean GI of HRT nonusers
Study group: mean GI of intervention subjects – mean GI of control subjects
4.2.3 Correlations between Diet GI and Nutrients

Table 4.5 summarizes the results of the correlation analysis to assess the relationships between diet GI and energy, macronutrients and various types of carbohydrates. The most consistent associations, seen at both time points and in all subjects, were with specific types of carbohydrates. Starch was positively associated with diet GI ($r = 0.31$ to $0.34$, $p < 0.0001$), whereas total fibre and lactose were negatively associated with diet GI ($r = -0.24$ to $-0.36$, $p < 0.0001$). Fructose was also negatively associated with diet GI at baseline in cases ($r = -0.17$, $p = 0.01$) and controls ($r = -0.18$, $p < .0001$), but only in controls at post-randomization ($r = -0.28$, $p < 0.0001$). Sucrose tends to be positively associated with diet GI but the association was only significant with baseline diet GI in controls ($r = 0.11$, $p = 0.02$) and with post-randomization diet GI in cases ($r = 0.17$, $p = 0.02$). Total sugars were not associated with diet GI except in controls at post-randomization ($r = -0.25$, $p < .0001$) when fructose and lactose had a strong negative association and sucrose was not associated with diet GI.

The relationships of diet GI with energy and macronutrients were weaker and less consistent than with types of CHO, seen usually at one time point and in either cases or controls. Energy was positively associated with baseline diet GI in controls ($r = 0.14$, $p = 0.004$). Fat was positively associated in controls with diet GI at baseline ($r = 0.15$, $p = 0.001$) and post-randomization ($r = 0.15$, $p = 0.002$). Protein tends to be negatively correlated with diet GI but the relationship was only significant with baseline diet GI in cases ($r = -0.15$, $p = 0.03$) and with post-randomization diet GI in controls ($r = -0.10$, $p = 0.03$). Total CHO was positively associated with post-randomization diet GI in cases ($r = 0.14$, $p = 0.05$).

The correlations of overall diet GI with energy and nutrients were similar to the findings for post-randomization diet GI (data not shown).

I conducted linear regression analysis to examine which nutrients were independently associated with diet GI (Table 4.6). The first model included the nutrients with the most consistent relationship with diet GI: starch, fibre, lactose, fructose, sucrose and protein. With the types of carbohydrate in the analysis, protein was no longer associated with diet GI (coefficient in cases = -0.02, $p = 0.18$; coefficient in controls = 0.007, $p = 0.40$) and was excluded from the model. Table 4.6 shows the results of the linear regression analysis with the five carbohydrates and diet GI at baseline and post-randomization. These variables were independently associated with baseline diet GI in cases and controls, except fructose which
did not attain significance in cases. The five predictor variables explain 39% to 49% of variation in baseline diet GI ($r^2$ controls = 0.39, $p < 0.0001$ and $r^2$ cases = 0.49, $p < 0.0001$).

The standardized regression coefficients suggest that total fibre (coefficient in cases = -0.58 and coefficient in controls = -0.46) and starch (coefficient in cases = 0.64 and coefficient in controls = 0.49) have the strongest effect on baseline diet GI, and fructose has the smallest effect (coefficient in cases = -0.08 and coefficient in controls = -0.09) (data not shown). Therefore, when fructose was removed from the model, the proportion of diet GI explained by the remaining carbohydrates only decreased by 1% and the effects of the other carbohydrates were unchanged.

Energy was added to the model shown in Table 4.6, but was not associated with baseline diet GI (coefficient in cases = 0.00, $p = 0.89$; coefficient in controls = 0.001, $p = 0.16$). Then, fat was added to the model, but it was also not associated with baseline diet GI (coefficient in cases = 0.005, $p = 0.60$; coefficient in controls = 0.01, $p = 0.10$).

The results of the linear regression analysis with post-randomization diet GI in both cases and controls were similar to baseline diet GI. The five predictor variables explain 50% to 53% of variation in post-randomization diet GI ($r^2$ controls = 0.50, $p < 0.0001$ and $r^2$ cases = 0.53, $p < 0.0001$). Energy, protein and total fat were added to the model but they were also not associated with post-randomization diet GI (data not shown).

I also ran the models using overall mean diet GI and the results were very similar to the findings with post-randomization diet GI (data not shown).

In summary, although specific carbohydrates are correlated in the expected direction and are independently associated with diet GI, they only explain 50% of the variation in GI. Therefore, Diet GI can provide additional information about dietary intake that cannot be measured by nutrient intake.
### Table 4.5 Pearson correlations of diet GI with energy and nutrients

<table>
<thead>
<tr>
<th>Types of Carbohydrate</th>
<th>Baseline Diet GI</th>
<th>Post-randomization Diet GI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient (p value)</td>
<td>Correlation Coefficient (p value)</td>
</tr>
<tr>
<td><strong>Total Fibre</strong></td>
<td>$-0.32$ ($&lt;0.0001$)</td>
<td>$-0.24$ ($&lt;0.0001$)</td>
</tr>
<tr>
<td><strong>ACHO</strong></td>
<td>$0.13$ (0.05)</td>
<td>$0.12$ (0.01)</td>
</tr>
<tr>
<td><strong>Starch</strong></td>
<td>$0.33$ ($&lt;0.0001$)</td>
<td>$0.31$ ($&lt;0.0001$)</td>
</tr>
<tr>
<td><strong>Sucrose</strong></td>
<td>$0.12$ (0.07)</td>
<td>$0.11$ (0.02)</td>
</tr>
<tr>
<td><strong>Lactose</strong></td>
<td>$-0.24$ ($&lt;0.0001$)</td>
<td>$-0.27$ ($&lt;0.0001$)</td>
</tr>
<tr>
<td><strong>Fructose</strong></td>
<td>$-0.17$ (0.01)</td>
<td>$-0.18$ ($&lt;0.0001$)</td>
</tr>
<tr>
<td><strong>Total Sugars</strong></td>
<td>$-0.07$ (0.29)</td>
<td>$-0.09$ (0.07)</td>
</tr>
</tbody>
</table>

### Energy and Macronutrients

<table>
<thead>
<tr>
<th></th>
<th>Baseline Diet GI</th>
<th>Post-randomization Diet GI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient (p value)</td>
<td>Correlation Coefficient (p value)</td>
</tr>
<tr>
<td><strong>Energy</strong></td>
<td>$0.09$ (0.18)</td>
<td>$0.14$ (0.004)</td>
</tr>
<tr>
<td><strong>Total Fat</strong></td>
<td>$0.09$ (0.19)</td>
<td>$0.15$ (0.001)</td>
</tr>
<tr>
<td><strong>Protein</strong></td>
<td>$-0.15$ (0.03)</td>
<td>$-0.03$ (0.47)</td>
</tr>
<tr>
<td><strong>Total CHO</strong></td>
<td>$0.08$ (0.22)</td>
<td>$0.08$ (0.08)</td>
</tr>
</tbody>
</table>
Table 4.6 Regression coefficients of diet GI with types of carbohydrate

<table>
<thead>
<tr>
<th></th>
<th>Baseline Diet GI</th>
<th>Post-randomization Diet GI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td>Total Fibre</td>
<td>-0.25</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)</td>
<td>(&lt;0.0001)</td>
</tr>
<tr>
<td>Starch</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)</td>
<td>(&lt;0.0001)</td>
</tr>
<tr>
<td>Sucrose</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(&lt;0.0001)</td>
</tr>
<tr>
<td>Lactose</td>
<td>-0.10</td>
<td>-0.11</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)</td>
<td>(&lt;0.0001)</td>
</tr>
<tr>
<td>Fructose</td>
<td>-0.03</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(&lt;0.03)</td>
</tr>
<tr>
<td>R²</td>
<td>0.49</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)</td>
<td>(&lt;0.0001)</td>
</tr>
</tbody>
</table>
4.3 Association of Diet GI with Breast Cancer Risk - Conditional Logistic Regression

Conditional logistic regression was conducted to account for matching and adjusting for potential confounders. We examined the association of diet GI with overall risk of breast cancer, and risk of ER positive and ER negative breast cancer. First we conducted analysis with diet GI alone, then we conducted the analysis with adjustments for breast cancer risk factors: parity, menopausal status, hormone ever use, family history, smoking ever, age at menarche, age at first live birth, number of live births for parous women, weight and energy intake and these results are shown in Table 4.7. Finally we did the regression analysis with the types of CHO found to be associated with ER positive and ER negative breast cancer in the DBCPT (see figure 2.5) and these results are shown in Table 4.8. Including the CHO's in the model allowed us to determine if their effects were independent of diet GI. The analyses were conducted using diet GI at baseline, post-randomization and overall mean.

4.3.1 Univariate and Multivariate Models with Diet GI and Risk Factors for Breast Cancer

Baseline diet GI was not associated with overall risk of breast cancer before (p = 0.45) or after adjustment for risk factors (p = 0.65). These results persisted at post-randomization and with the overall mean of food records. Diet GI was not associated with overall risk of breast cancer.

Baseline diet GI was also not associated with risk of ER positive breast cancer before (p = 0.91) or after adjustment for risk factors (p = 0.65), and remained non-significant at post-randomization and with the overall mean of food records in the univariate and multivariate models.

Diet GI at baseline and post-randomization were also not associated with risk of ER negative breast cancer before and after adjustment for risk factors. However, diet GI based on the overall mean of food records was positively associated with risk of ER negative breast cancer before adjustment for risk factors, but the relationship was borderline significant (p = 0.07).
Table 4.7 Diet GI and breast cancer risk - conditional logistic regression

<table>
<thead>
<tr>
<th></th>
<th>All Breast Cancer</th>
<th>ER positive</th>
<th>ER Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>220 cases,</td>
<td>167 cases,</td>
<td>42 cases,</td>
</tr>
<tr>
<td></td>
<td>440 controls (^a)</td>
<td>334 controls (^b)</td>
<td>84 controls (^c)</td>
</tr>
<tr>
<td><strong>OR (95% CI)</strong></td>
<td><strong>P value</strong></td>
<td><strong>OR (95% CI)</strong></td>
<td><strong>P value</strong></td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI</td>
<td>1.02 (0.97, 1.08)</td>
<td>1.00 (0.94, 1.06)</td>
<td>1.08 (0.96, 1.22)</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors</td>
<td>1.01 (0.96, 1.07)</td>
<td>0.98 (0.92, 1.05)</td>
<td>1.09 (0.94, 1.25)</td>
</tr>
<tr>
<td><strong>Post-randomization</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI</td>
<td>1.03 (0.95, 1.11)</td>
<td>1.00 (0.92, 1.09)</td>
<td>1.09 (0.90, 1.32)</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors</td>
<td>1.01 (0.93, 1.10)</td>
<td>0.99 (0.90, 1.08)</td>
<td>1.13 (0.89, 1.43)</td>
</tr>
<tr>
<td><strong>Overall Mean</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI</td>
<td>1.05 (0.97, 1.13)</td>
<td>1.00 (0.93, 1.10)</td>
<td>1.19 (0.98, 1.44)</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors</td>
<td>1.03 (0.95, 1.12)</td>
<td>0.99 (0.90, 1.09)</td>
<td>1.19 (0.93, 1.53)</td>
</tr>
</tbody>
</table>

**Legend:**

\(^a\) Data missing at post-randomization: 13 cases, 10 controls

\(^b\) Data missing at post-randomization: 10 cases, 8 controls

\(^c\) Data missing at post-randomization: 3 cases, 2 controls
4.3.2 Multivariate Analysis with Diet GI, Risk Factors and Type of CHO

Table 4.8 summarizes the results of the conditional logistic regression analysis for risk of ER positive breast cancer with starch in the model (top) and ER negative breast cancer with total sugars in the model (bottom).

The results from the DBCPT indicated that starch was negatively associated with ER positive breast cancer (see figure 2.5). When starch was added to the model with risk factors and diet GI, starch was significantly associated with risk of ER positive breast cancer at baseline (OR 0.98, 95% CI 0.97, 0.99, p = 0.0002), post randomization (OR 0.99, 95% CI 0.97, 1.0, p = 0.04) and with the mean of all food records (OR 0.99, 95% CI: 0.97, 1.00, p = 0.03) but diet GI remained non-significant at all time points (baseline p = 0.72, post-randomization p = 0.58, overall mean p = 0.52). To compare to the results for starch seen in figure 2.5 that showed the OR across the IQR, our findings would suggest that with diet GI in the model, starch reduced the odds of developing ER positive breast cancer by 29% (IQOR 0.72, 95% CI: 0.52, 0.99) which is very similar to the initial findings of the DBCPT (IQOR 0.73, 95% CI: 0.55, 0.98). In a model with mutual adjustment, starch remained negatively associated and diet GI was not associated with risk of ER positive breast cancer.

As total sugars were positively associated with risk of ER negative cancer in the DBCPT analysis, total sugars were added to the model with risk factors and diet GI. Total sugars were significantly associated with risk of ER negative breast cancer at all time points (baseline OR 1.03, 95% CI 1.01, 1.06, p = 0.01), (post-randomization OR 1.04, 95% CI 1.02, 1.08, p = 0.003) (overall mean OR 1.05, 95% CI: 1.02, 1.08, p = 0.002). With total sugars added to the model, diet GI became significantly associated with risk of ER negative breast cancer at baseline (OR 1.23, 95% CI: 1.01, 1.50, p = 0.04). Similar results were seen at post-randomization (OR 1.38, 95% CI 1.00, 1.91, p = 0.05) and with the overall mean of all food records (OR 1.48, 95% CI: 1.06, 2.08). The OR across the IQR (Diet GI IQR = 3) suggests that a higher GI diet increases the odds 2.7 times of developing ER-negative breast cancer compared to controls (IQOR 2.7, 95% CI: 1.01, 6.97). Similarly, a high intake of total sugars increases the odds 3.9 times of developing ER-breast cancer compared to controls.
Table 4.8 Diet GI and breast cancer risk by estrogen receptor status

<table>
<thead>
<tr>
<th></th>
<th><strong>ER Positive</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>167 cases, 334 controls a</td>
<td></td>
</tr>
<tr>
<td><strong>OR (95% CI)</strong></td>
<td><strong>P value</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors and starch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI</td>
<td>1.01 (0.94, 1.09)</td>
<td>0.72</td>
</tr>
<tr>
<td>Starch</td>
<td>0.98 (0.97, 0.99)</td>
<td><strong>0.0002</strong></td>
</tr>
<tr>
<td><strong>Post-randomization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors and starch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI</td>
<td>1.03 (0.93, 1.14)</td>
<td>0.58</td>
</tr>
<tr>
<td>Starch</td>
<td>0.99 (0.97, 1.00)</td>
<td><strong>0.04</strong></td>
</tr>
<tr>
<td><strong>Overall Mean</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors and starch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI</td>
<td>1.03 (0.93, 1.14)</td>
<td>0.52</td>
</tr>
<tr>
<td>Starch</td>
<td>0.99 (0.97, 1.00)</td>
<td><strong>0.03</strong></td>
</tr>
</tbody>
</table>

|                      | **ER Negative** |                  |
|                      | 42 cases, 84 controls b |                  |
| **Baseline**         |                  |                  |
| Diet GI adjusted for risk factors and total sugars |                  |                  |
| Diet GI              | 1.23 (1.01, 1.50) | **0.04**         |
| Total Sugars         | 1.03 (1.01, 1.06) | **0.01**         |
| **Post-randomization** |                  |                  |
| Diet GI adjusted for risk factors and total sugars |                  |                  |
| Diet GI              | 1.38 (1.00, 1.91) | **0.05**         |
| Total Sugars         | 1.04 (1.02, 1.08) | **0.003**        |
| **Overall Mean**     |                  |                  |
| Diet GI adjusted for risk factors and total sugars |                  |                  |
| Diet GI              | 1.48 (1.06, 2.08) | **0.02**         |
| Total Sugars         | 1.05 (1.02, 1.08) | **0.002**        |

Legend:
- Data missing at post-randomization: 10 cases, 8 controls
- Data missing at post-randomization: 3 cases, 2 controls
(IQOR 3.9, 95% CI 1.57, 9.47) which is a stronger association than seen in the DBCPT
(IQOR 2.7, 95% CI: 1.3, 5.64). There was an independent, positive association between total
sugars and risk of ER negative breast cancer. Diet GI was positively associated with risk of
ER negative breast cancer when the analysis controlled for risk factors and total sugars.

4.4 Diet GI and Breast Cancer Phenotype in Cases-only Analysis

I will present the results from three types of analysis using data from breast cancer
cases only to examine the association between diet GI and: (1) risk of developing hormone
receptor negative versus hormone receptor positive breast cancer (2) tumor size; and (3) risk
of node negative versus node positive breast cancer.

4.4.1 Risk of Hormone Receptor Negative versus Hormone Receptor Positive Breast Cancer

4.4.1.1 Characteristics of Subjects by Hormone Receptor status

The distribution of risk factors among cases with ER positive and ER negative breast
cancer at the time of entry to the DBCPT is shown in Table 4.9. The mean weight and BMI
for ER positive cases (mean weight 63.9, SD 8.4; mean BMI 23.9, SD 2.7) were higher than
ER negative cases (mean weight 60.8, SD 9.0; mean BMI 22.8, SD 2.3). These differences
were tested using an independent group t test, and were shown to be significant (p (weight) =
0.05; p (BMI) = 0.01). Age, height, age at menarche, parity, age at first live birth, number of
children for parous women, family history, menopausal status and hormone ever use were
not significantly different between ER positive and ER negative cases.

This distribution of risk factors among cases with PR positive and PR negative, as
well as among cases with concordant ER/PR receptors was very similar to the results
described above (data not shown).
Table 4.9  Selected baseline characteristics of breast cancer cases

<table>
<thead>
<tr>
<th></th>
<th>ER Positive (n = 167)</th>
<th>ER Negative (n = 42)</th>
<th>P-value(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD) or %</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Group (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>50.3</td>
<td>59.5</td>
<td>.29</td>
</tr>
<tr>
<td>Control</td>
<td>49.7</td>
<td>40.5</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>48.6 (6.2)</td>
<td>47.9 (6.9)</td>
<td>.54</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>63.9 (8.4)</td>
<td>60.8 (9.0)</td>
<td>.05</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>163.6 (6.4)</td>
<td>162.9 (6.9)</td>
<td>.57</td>
</tr>
<tr>
<td>BMI (kg/m(^2))</td>
<td>23.9 (2.7)</td>
<td>22.8 (2.3)</td>
<td>.01</td>
</tr>
<tr>
<td>Age at Menarche (years)</td>
<td>12.9 (1.4)</td>
<td>12.7 (1.3)</td>
<td>.35</td>
</tr>
<tr>
<td>Parity (% parous)</td>
<td>67.7</td>
<td>64.3</td>
<td>.68</td>
</tr>
<tr>
<td>Age at First Live Birth (years)</td>
<td>26.3 (4.4)</td>
<td>25.6 (3.7)</td>
<td>.34</td>
</tr>
<tr>
<td>Number of Children for Parous Women</td>
<td>2.2 (0.6)</td>
<td>2.1 (0.9)</td>
<td>.92</td>
</tr>
<tr>
<td>First Degree Relative with Breast Cancer (% yes)</td>
<td>23.4</td>
<td>26.2</td>
<td>.70</td>
</tr>
<tr>
<td>Menopausal Status (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premenopausal</td>
<td>70.7</td>
<td>76.2</td>
<td>.48</td>
</tr>
<tr>
<td>Postmenopausal</td>
<td>29.3</td>
<td>23.8</td>
<td></td>
</tr>
<tr>
<td>Hormone Use (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28.7</td>
<td>38.1</td>
<td>.24</td>
</tr>
<tr>
<td>No</td>
<td>71.3</td>
<td>61.9</td>
<td></td>
</tr>
</tbody>
</table>

**Legend:**

\(^a\) P value from 2 sample t-test for: age, weight, height, BMI, age at menarche, age at first live birth; Wilcoxon Mann-Whitney test for number of children born to parous women; Chi square test for categorical variables: study group, parity, family history, menopausal and HRT ever use.

Data missing for ER status for 11 breast cancer cases.
4.4.1.2 Association of Diet GI with Breast Cancer Risk by Hormone Receptor Status

The approach to the logistic regression analysis was similar to the strategy used in the stratified analysis in section 4.3. The first step was univariate analysis with diet GI alone, then we added breast cancer risk factors to the model and finally we included total sugars in the model because the effect of diet GI on the risk of ER negative breast cancer was dependent on total sugars (section 4.3.2). We considered estrogen receptors status, progesterone receptors status and then examined risk for hormone receptor negative and positive sub-types of breast cancer.

A. Risk of ER negative versus ER positive breast cancer

We examined the association of diet GI with risk of developing ER negative versus ER positive breast cancer based on data from 42 ER negative and 167 ER positive cases (Table 4.10 left column). In the unadjusted analysis, diet GI was not associated with risk of developing ER negative compared to ER positive breast cancer at baseline (p = 0.18), post-randomization (p = 0.38) or with the mean of all food records (p = 0.13). After adjustment for breast cancer risk factors, diet GI remained non-significant at all time points (baseline p = 0.19) (post-randomization p = 0.22) (overall mean p = 0.11). Adding total sugars to the model the strengthened the association between diet GI and risk of ER negative breast cancer slightly. Diet GI was positively associated with the risk of developing ER negative versus ER positive breast cancer, but the association was not significant (overall mean OR 1.16, 95% CI: 0.98, 1.38, p = 0.09).

B. Risk of PR negative versus PR positive breast cancer

We examined the association of diet GI with the risk of developing PR negative versus PR positive breast cancer based on data from 67 PR negative and 123 PR positive cases (Table 4.10 middle column). In the univariate analysis, diet GI was not associated with risk of developing PR negative compared to PR positive breast cancer at baseline (p = 0.30), post-randomization (p = 0.15) or with the mean of all food records (p = 0.12). After adjustment for breast cancer risk factors, diet GI remained non-significant at baseline (p = 0.16) and post-randomization (OR 1.14, 95% CI 0.99, 1.31, p = 0.07) but was found to be significantly associated with risk of PR negative breast cancer when the mean of all food
records was used to determine the diet GI (OR 1.17, 95% CI: 1.00, 1.36, p = 0.05). When total sugars were added to the model, the results were similar to the previous analysis with adjustments breast cancer risk factors.

C. Risk of ER-/PR- versus ER+/PR+ breast cancer

We examined the association of diet GI with the risk of developing hormone receptor negative versus hormone receptor positive breast cancer based on data from 33 ER-/PR- and 116 ER+/PR+ cases (Table 4.10 right column). In the univariate analysis or in a model with risk factors and total sugars, baseline diet GI was positively associated with risk of hormone receptor breast cancer but the association is not significant (p = 0.08 to 0.11). At post-randomization, the univariate analysis was not significant (p = 0.15) but with the addition of risk factors to the model, the positive relationship between diet GI and risk of hormone receptor negative breast cancer became significant (OR 1.25, 95% CI: 1.02, 1.52, p = 0.03). The strength of this association was similar when total sugars were added to the model (OR 1.29, 95% CI: 1.05, 1.59, p = 0.02). With all food records used to estimate the diet GI we saw a significant positive association in the univariate analysis (OR 1.18, 95% CI: 1.00, 1.40, p = 0.05). The addition of risk factors strengthened the risk estimate associated with diet GI (OR 1.32, 95% CI: 1.06, 1.65, p = 0.01). Adding total sugars in the model had very little effect (OR 1.37, 95% CI: 1.09, 1.72, p = 0.008). The OR across the IQR suggests that when the diet GI increases by 3 points, the odds are 2.3 times higher of developing ER-/PR-breast cancer compared to ER+/PR+ breast cancer (OR 2.3, 95% CI 1.19, 4.49).
Table 4.10  Diet GI with risk of hormone receptor negative versus positive breast cancer

<table>
<thead>
<tr>
<th></th>
<th>ER- vs ER+</th>
<th>PR- vs PR+</th>
<th>ER-/PR- vs ER+/PR+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>42 ER-, 167 ER+</td>
<td>67 PR-, 123 PR+</td>
<td>33 ER-/PR -, 116 ER+/PR+</td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td>OR (95% CI)</td>
<td>P value</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>1.07 (0.99, 1.19)</td>
<td>0.18</td>
<td>1.05 (0.96, 1.15)</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors d</td>
<td>1.08 (0.96, 1.22)</td>
<td>0.19</td>
<td>1.08 (0.97, 1.19)</td>
</tr>
<tr>
<td>Diet GI, adjusted for risk factors and total sugars</td>
<td>1.08 (0.96, 1.22)</td>
<td>0.18</td>
<td>1.10 (0.99, 1.23)</td>
</tr>
<tr>
<td><strong>Post-randomization</strong></td>
<td>OR (95% CI)</td>
<td>P value</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>1.07 (0.93, 1.23)</td>
<td>0.38</td>
<td>1.10 (0.97, 1.13)</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors d</td>
<td>1.10 (0.94, 1.29)</td>
<td>0.22</td>
<td>1.14 (0.99, 1.31)</td>
</tr>
<tr>
<td>Diet GI, adjusted for risk factors and total sugars</td>
<td>1.12 (0.95, 1.31)</td>
<td>0.17</td>
<td>1.15 (1.00, 1.32)</td>
</tr>
<tr>
<td><strong>Overall Mean</strong></td>
<td>OR (95% CI)</td>
<td>P value</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>1.11 (0.97, 1.28)</td>
<td>0.13</td>
<td>1.11 (0.98, 1.25)</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors e</td>
<td>1.15 (0.97, 1.36)</td>
<td>0.11</td>
<td>1.17 (1.00, 1.36)</td>
</tr>
<tr>
<td>Diet GI, adjusted for risk factors and total sugars</td>
<td>1.16 (0.98, 1.38)</td>
<td>0.09</td>
<td>1.17 (1.01, 1.37)</td>
</tr>
</tbody>
</table>
**Legend:**

a  39 ER-, 157 ER+ for post randomization
b  63 PR-, 117 PR+ for post randomization
c  31 ER-/PR-, 111 ER+/PR+ for post randomization
d  Risk factors are: baseline age, age at menarche, parity, age at first birth, number of live births, menopausal status, use of hormone replacement therapy (ever/never), first degree relative with breast cancer, smoking (ever/never), baseline energy intake and baseline weight;
e  Same risk factors as above but mean of post-randomization energy intake and weight included instead of baseline values
4.4.2 Association of Diet GI with Tumor Size

The histograms in Figure 4.3 show that the distribution of data on tumor size from 207 subjects had a positive skew but after log transformation the data became more normally distributed.

The association of diet GI with log tumor size was determined using a general linear model (Table 4.11). The approach to the analysis was similar to the strategy used in the stratified analysis in section 4.3 and the hormone receptor status analysis in section 4.4.1. The first step was univariate analysis with diet GI alone, then we added breast cancer risk factors to the model and finally we included total sugars in the model. In the unadjusted analysis, diet GI was not associated with log tumor size at baseline (p = 0.14), post-randomization (p = 0.34) or with the mean of all food records (p = 0.22). After adjustment for breast cancer risk factors, diet GI remained non-significant at all time points (baseline p = 0.34) (post-randomization p = 0.55) (overall mean p = 0.43). When total sugars were added to the model the results were unchanged at all time points. Diet GI was not associated with tumor size.

4.4.3 Association of Diet GI with risk of Nodal Status

The association of diet GI on nodal status was determined using logistic regression (Table 4.12). The approach to the analysis was similar to the strategy used for other phenotype analyses with the addition of a step to add log tumor size to the model because it is strongly associated with risk of nodal involvement (coefficient -1.31, SE 0.32, p < 0.0001).

In the unadjusted analysis, diet GI was not associated with nodal status at baseline (p = 0.96), post-randomization (p = 0.57) or with the mean of all food records (p = 0.61). After adjustments for log tumor size, breast cancer risk factors and total sugars, diet GI remained non-significant at all time points. Diet GI was not associated with nodal status (p = 0.73 - 0.98).

All the analyses in section 4.4 were also conducted with randomized group in the model and this did not affect the results.
Figure 4.3 Histograms of tumor size data

N = 207
Mean = 1.74
SD = 1.29
Median = 1.50

N = 207
Mean = 0.35
SD = 0.67
Median = 0.41
Table 4.11 Association of diet GI with log tumor size (log cm)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>SE</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>0.02</td>
<td>0.01</td>
<td>0.14</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.01</td>
<td>0.01</td>
<td>0.34</td>
</tr>
<tr>
<td>Diet GI, adjusted for risk factors and total sugars</td>
<td>0.02</td>
<td>0.01</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>Post-randomization</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>0.02</td>
<td>0.02</td>
<td>0.34</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.01</td>
<td>0.02</td>
<td>0.55</td>
</tr>
<tr>
<td>Diet GI, adjusted for risk factors and total sugars</td>
<td>0.01</td>
<td>0.02</td>
<td>0.53</td>
</tr>
<tr>
<td><strong>Overall Mean</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>0.02</td>
<td>0.02</td>
<td>0.22</td>
</tr>
<tr>
<td>Diet GI adjusted for risk factors&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.02</td>
<td>0.02</td>
<td>0.43</td>
</tr>
<tr>
<td>Diet GI, adjusted for risk factors and total sugars</td>
<td>0.02</td>
<td>0.02</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Legend:
<sup>a</sup> Baseline: n = 207  Post-Randomization: n = 196
<sup>b</sup> risk factors are: baseline age, age at menarche, parity, age at first birth, number of live births, menopausal status, use of hormone replacement therapy (ever/never), first degree relative with breast cancer, smoking (ever/never), baseline energy intake and baseline weight;
<sup>c</sup> same risk factors as above but mean of post-randomization energy intake and weight included instead of baseline values
Table 4.12 Diet GI with risk of nodal involvement

<table>
<thead>
<tr>
<th></th>
<th>Node Negative versus</th>
<th>Node Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>145 Node -, 61 Node +</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>1.00 (0.91, 1.09)</td>
<td>0.96</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size</td>
<td>1.02 (0.92, 1.12)</td>
<td>0.73</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size and risk factors  b</td>
<td>1.03 (0.93, 1.14)</td>
<td>0.60</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size, risk factors and total sugars</td>
<td>1.02 (0.92, 1.13)</td>
<td>0.65</td>
</tr>
<tr>
<td><strong>Post-randomization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>0.97 (0.86, 1.09)</td>
<td>0.57</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size</td>
<td>0.99 (0.88, 1.13)</td>
<td>0.92</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size and risk factors  b</td>
<td>1.00 (0.88, 1.15)</td>
<td>0.98</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size, risk factors and total sugars</td>
<td>1.00 (0.88, 1.16)</td>
<td>0.90</td>
</tr>
<tr>
<td><strong>Overall Mean</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet GI alone</td>
<td>0.97 (0.86, 1.09)</td>
<td>0.61</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size</td>
<td>1.00 (0.88, 1.13)</td>
<td>0.96</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size and risk factors  c</td>
<td>1.01 (0.88, 1.16)</td>
<td>0.91</td>
</tr>
<tr>
<td>Diet GI adjusted for log tumor size, risk factors and total sugars</td>
<td>1.02 (0.88, 1.17)</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Legend:

<table>
<thead>
<tr>
<th>a 138 node negative, 57 node positive for post randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>b risk factors are: baseline age, age at menarche, parity, age at first birth, number of live births, menopausal status, use of hormone replacement therapy (ever/never), first degree relative with breast cancer, smoking (ever/never), baseline energy intake and baseline weight;</td>
</tr>
<tr>
<td>c same risk factors as above but mean of post-randomization energy intake and weight included instead of baseline values</td>
</tr>
</tbody>
</table>

4.5 Summary of Main Results

Breast Cancer Risk (overall analysis and stratified by hormone receptor status)

1. Diet GI was not associated with overall breast cancer risk, alone or after adjustment for risk factors.
2. Diet GI was not associated with risk of ER positive breast cancer, alone or after adjustment for risk factors and starch intake.
3. Diet GI was positively associated with risk of ER negative breast cancer after adjustment for risk factors and intake of total sugars in the analysis stratified by hormone receptor status.
4. The inverse association of starch intake with risk of ER positive breast cancer was not altered or explained by diet GI.
5. The positive association of total sugars with risk of ER negative breast cancer was strengthened by diet GI.

Breast Cancer Phenotype (case-only analysis)

1. Diet GI was positively associated with the risk of hormone negative breast cancer and this relationship was statistically significant at post-randomization after adjustment for risk factors, and with diet GI alone when we used the mean of all food records to estimate diet GI. These results are based on a small number of cases and secondary analysis and should be viewed with caution.
2. Diet GI was not associated with tumor size or nodal status.
5.1 Overview of results

I conducted a case-control study nested in a cohort of subjects from the DBCPT to investigate the association of diet GI on breast cancer risk and phenotype using dietary data from food records. I did not see an association of diet GI with overall breast cancer risk. This finding was similar to the individual results from most observational studies (79-85;87;113-117) and the meta-analysis conducted in 2008 (119).

I examined the association of diet GI with breast cancer risk by hormone receptor status of the tumor. I found no association with ER positive breast cancer, but a significant association between diet GI and ER negative breast cancer. In the univariate analysis and after adjustment for breast cancer risk factors, diet GI was positively associated with risk of ER negative breast cancer but the relationship was borderline significant. We added total sugars to the model for two reasons: (1) We had previously seen they were associated with risk of ER negative breast cancer; and (2) Total sugars were related to diet GI. When total sugars were included as a covariate, the association between diet GI and ER negative breast cancer risk became stronger and statistically significant (IQOR 2.65, 95% CI: 1.01, 6.97). Including GI in the model strengthened the association of total sugars with breast cancer risk seen in the DBCPT. Therefore, GI and total sugars appeared to have strong, independent effects on risk of ER negative breast cancer.

I also investigated the association of diet GI with risk among breast cancer cases, of developing hormone negative versus hormone positive breast cancer and found a significant positive association of diet GI with risk of developing hormone negative breast cancer. Diet GI was associated with an increased risk of ER negative breast cancer but the association was not significant. Diet GI was positively associated with risk of PR negative breast cancer, but the strongest association of diet GI with hormone receptor negative breast cancer was seen in analysis using both hormone receptors. In the univariate analysis diet GI at baseline and post-randomization were associated with an increased risk of ER negative/PR negative breast cancer, but the association became significant when we used the overall mean of all food records. The risk estimate increased with the addition of other risk factors to the model, and increased slightly again with adjustment for total sugars. In the case only analysis, diet GI
appeared to have a strong association with the risk of developing ER negative/PR negative breast cancer and the effect was evident without adjustment for total sugars.

The role of total sugars in the determination of the association between diet GI and risk of ER negative breast cancer in the stratified analysis may be related to the impact of total sugars on the diet of ER negative controls. I examined case control differences in the amount of total sugars consumed and the correlation between total sugars with diet GI. ER negative controls consumed significantly less total sugars than ER negative cases; and less than other controls matched to ER positive cases but the difference was not significant. There was no significant difference in the amount of total sugars consumed by ER positive and ER negative cases. Correlation analysis showed that total sugars were significantly correlated with diet GI in controls but not in cases. Therefore, total sugars may have cofounded the relationship between diet GI and risk of ER negative breast cancer in the stratified analysis and adjustment for total sugars was necessary to see the risk associated with diet GI. As total sugars were not correlated with diet GI in cases, their addition to the model in the case only analysis had little effect.

Few studies have examined the effect of GI on breast cancer risk by hormone receptor status (79;81-83). Our findings are consistent with two studies that observed a significant positive association of GI with hormone receptor negative cancers (79;83). Lajous et al did not see association with GI, but they found total CHO and glycemic load were positively associated with ER negative tumors (81). The Australian study by Giles et al had fewer cases than the other studies which limited statistical power to detect an association (82).

Risk factors for breast cancer may differ by hormone receptor sub-types. Epidemiological studies of breast cancer risk factors and hormone receptor status found some reproductive factors as well as BMI and adult weight gain are associated with increased risk of ER positive, but not ER negative breast cancer. There has been some inconsistency among studies of alcohol intake but large, recent studies with high quality data about alcohol intake suggest a strong association between alcohol and ER positive breast cancer risk, and little, if any association with risk ER negative tumors (54). There is limited information and conflicting results from the few studies of fat and/or CHO intake to identify an association with these specific nutrients and hormone receptors. Secondary analysis of the WHI found the intervention was inversely associated with ER positive/PR negative breast cancer (HR
0.64 95% CI: 0.49-0.84), but not other receptor sub-types (101). Three cohort studies of dietary fat and ER status found no association (145-147). The Iowa Women’s Health Study observed an inverse association of total CHO with ER positive/PR positive breast cancer (147). The NIH-AARP Study found a negative association with dietary fibre and ER negative/PR negative breast cancer (95).

In our study we found a significant inverse association between starch and ER positive breast cancer. This observation is consistent with our previous results from the DBCPT, the Iowa Women’s Health Study and the ecological studies of per capita starch consumption and breast cancer mortality rate (72). Furthermore, including GI in the model did not alter or explain the relationship between starch and ER positive breast cancer. Glycemic load is a measure of the combined effect of CHO quantity and quality and is defined as the food GI multiplied by the amount of available CHO (g) divided by 100. Three studies found a significant positive association between glycemic load, but not GI, and breast cancer risk (80;81;116). The Shanghai Women’s Health Study found dietary glycemic load and total CHO intake were associated with breast cancer risk in premenopausal women (80). The French cohort of the EPIC study found glycemic load and total CHO intake were associated with risk of ER negative breast cancer (81). A Mexican case-control study found glycemic load was associated with breast cancer risk overall and particularly in post-menopausal women (116). In our study, total CHO was very highly correlated with glycemic load ($r^2 = 0.95$) and since most studies saw similar effects with CHO intake as with glycemic load, it seems unlikely that glycemic load would provide new information to our study. As it cannot be ruled out, I intend to further investigate the association of glycemic load and breast cancer risk and phenotype in the future.

Within the last decade, epidemiological studies have been investigating the role of dietary patterns in the development of breast cancer as an alternative to the single nutrient approach. Dietary patterns are a qualitative measure that describes groups of foods that contribute to a particular style of eating. The most common patterns studied are called “health” or “prudent” and tend to include fruit, vegetables, whole grains, fish, poultry and low fat dairy products but not sweets and potatoes. Six large cohort studies have explored associations of these dietary patterns with breast cancer risk by hormone receptor status (148-153). Four out of six studies found a protective effect with hormone receptor negative breast cancer
cancer, three with ER negative (149;152;153), and one with ER positive/PR negative breast cancer (150). No study found the healthy/prudent patterns were associated with ER positive breast cancer. These results from dietary patterns tend to be supportive of our work because one could hypothesize that the health/prudent patterns could have a low GI value due to the predominance of low GI foods.

To our knowledge this is the first study to examine the effect of GI on tumor size or nodal status. Our results indicate that GI was not related to tumor size or nodal status.

5.2 Potential Biological Mechanisms

We postulated initially that a high GI diet could increase breast cancer risk by increasing postprandial insulin levels. Hyperinsulinemia may play a role in the etiology of breast cancer directly or indirectly through its association with obesity, diabetes, and growth factors, such as IGF1. A pooled analysis of 17 prospective studies of IGF-1 and breast cancer risk found a positive association between IGF-1 and breast cancer risk in pre- and postmenopausal women but it was confined to estrogen receptor positive tumors (135). The Nurses’ Health Study found a significant association between type 2 diabetes and breast cancer risk among postmenopausal women with ER positive tumors, but not ER negative tumors (154). Since there has been no evidence linking insulin or IGF-1 to risk of ER negative tumors, it is unlikely that these factors explain the relationship we observed.

High GI diets have also been associated with HDL-C, triglycerides, adiponectin and markers of oxidative stress. All of these effects may also be associated with breast cancer risk and suggest mechanisms that may link high GI diets to increased risk of ER negative breast cancer are described below.

Some studies have shown that high GI diets are associated with lower serum HDL-C (155) (111;156;157) and higher triglyceride concentrations (111;158) and these lipids may be related to breast cancer risk. Many epidemiological studies have shown that HDL-C is inversely associated with breast cancer risk (159) but only one case-control study has examined lipid profiles and breast cancer risk by hormone receptor status (160). The study found high HDL-C levels were associated with lower risk of ER positive and ER negative breast cancer, but dyslipidemia, characterized by low HDL-C and elevated triglycerides was only associated with ER negative breast cancer. We have measured serum lipids for the
subjects in our nested case-control study and will assess if they are related to GI and risk of breast cancer.

There has been one cross sectional study of GI and adiponectin levels in women (161). In a sample of 902 women with type 2 diabetes from the Nurses’ Health Study, GI was inversely associated with plasma adiponectin. A similar study in diabetic men found the same result (162). Adiponectin is a protein secreted by adipose tissue and blood levels are inversely associated with body weight and BMI, and positively associated with age and HDL-C levels (163;164). Adiponectin receptors are expressed in human breast cancer cells, and mediate anti-proliferative and pro-apoptotic responses (165) (166). The anti-proliferative effects of adiponectin and breast cancer cell lines may induce growth arrest and apoptosis (167;168).

Six case-control studies and one cohort study have examined the association between adiponectin and risk of breast cancer. Most case-control studies, showed a significant association between lower levels of adiponectin and increased risk of breast cancer overall (164;169-173) and increased risk of postmenopausal breast cancer (169;170;174). The cohort study, by far the largest of the studies published to date, showed no overall association between levels of adiponectin and breast cancer risk, but found that lower adiponectin levels were associated with increased breast cancer risk in postmenopausal women who had never used hormone therapy and post-menopausal women with low circulating estradiol levels (173). Four studies, one cohort and three case-control, examined the association of adiponectin levels with breast cancer risk by ER status but the results are inconsistent. The cohort study found no association between adiponectin and risk of either ER positive (n = 399 cases) or ER negative breast cancer (n = 95 cases) (173). A Taiwanese study compared the adiponectin levels in cases with ER positive breast cancer (n = 60) with those in cases of ER negative breast cancer (n = 28) and found that adiponectin levels were not associated with ER status (172). Another Taiwanese study examined the association of low (<15.02 µg/mL) and high (>15.02 µg/mL) adiponectin levels with risk of breast cancer by ER status and found an inverse association of adiponectin level with risk ER positive breast cancer (n = 140 cases) but no association with ER negative breast cancer (n = 55 cases)(174). A Japanese study found the opposite result that low levels of adiponectin (<6.9 µg/mL) were associated with an increased risk of ER negative (n = 44 cases) but not ER positive breast cancer (n = 50
cases) (169). Therefore, if high GI diets reduce adiponectin levels, there is some evidence that lower levels of adiponectin may increase risk of breast cancer.

Oxidative stress occurs when the production of reactive oxygen species (ROS) exceeds the human antioxidant defenses and damage cellular DNA, protein and lipid molecules. Damage to DNA can lead to mutagenesis and increased risk of cancer. There is some evidence that GI may be related to oxidative stress. One cross-sectional study of 292 subjects found that diet GI was positively associated with plasma malonaldehyde (MDA), a marker of lipid peroxidation, and the relationship was stronger in subjects with a BMI < 26.5 kg/m² (175). In a small feeding study of 12 overweight or obese men, they found that after seven days on a low GI diet that total antioxidant capacity (oxygen radical absorbance capacity) was higher, but there was no change in urinary F2-isoprostane level, another marker of lipid peroxidation thought to be more accurate than MDA (176).

Direct evidence for the effect of oxidative stress on breast cancer development is limited to case-control studies and one nested case-control study from the Shanghai Women’s Health Study cohort. Several small case-control studies have shown that the concentration of plasma MDA was elevated in breast cancer patients compared to controls (177). The nested case-control study measured urinary F2-isoprostanes and found no difference between cases and controls. However, they observed different associations of isoprostanes with breast cancer risk depending on level of adiposity. Among obese women, urinary levels of isoprostanes were positively associated with breast cancer risk, but among women with BMI < 23 kg/m², they were negatively associated with breast cancer risk. Although this study used a more accurate measure of oxidative stress, it is difficult to interpret because the levels of urinary isoprostanes in Chinese women were much higher than levels normally found in healthy subjects in the United States (178). No studies examined the association of oxidative stress and breast cancer risk by ER status.

Although the connections between GI and oxidative stress, and even oxidative stress and breast cancer need to be confirmed, indirect evidence from the protective effects of “healthy diet patterns” suggest that diets rich in fruits, vegetables and other phytochemicals may be low in GI, rich in antioxidants and reduce breast cancer risk.
5.3 **Strengths and Limitations**

5.3.1 **Study Design**

We had a strong design for our study which minimized statistical biases. The case-control study nested within the cohort from the DBCPT ensured that data on breast cancer risk factors and diet was collected prospectively, before the development of cancer, which avoided recall bias. We strengthened the internal validity of our results as both cases and controls came from the same cohort.

The subjects from the DBCPT had extensive mammographic density which is a strong risk factor for breast cancer. Therefore, our results may not be representative of the general population. Mammographic density however, is not associated with ER status of the tumor (179-181) and the proportions of ER positive and ER negative tumours, 76% and 19% respectively, in our sample is similar to other studies of North American and European women. Therefore, it is unlikely that our finding with ER negative breast cancer was influenced by mammographic density. Previously in the trial, we found positive associations between weight and intake of animal protein with ER positive breast cancer which are consistent with other studies of women not selected on the basis of mammographic density. These observations suggest that although we studied a high risk group, our results on dietary factors and breast cancer risk and phenotype could indeed be generalizable.

Most of our subjects were lean, evident by the mean BMI for our population, 23.5 (SD 2.5). An association of GI and breast cancer may be detected more easily in women with a BMI < 25 kg/m². Four studies conducted analysis stratified by BMI and most found a significant positive association between GI and breast cancer risk for women with BMI < 25 kg/m² (84;85;112). Perhaps, the relationship of GI with breast cancer risk is easier to detect in leaner subjects and in a sub-type of breast cancer that is not related to obesity.

We obtained pathology reports with detailed information on the phenotype of breast cancer tumors for most of our cases which enabled us to examine the association between diet GI and hormone receptor status, tumor size and nodal status. Few studies analyzed results according to hormone receptor status, and none considered the effect of GI on tumour size and nodal status. Relying on pathology reports from various sources may have led to some misclassification of hormone receptor status due to inter-laboratory variation in their
The more significant limitations to our analysis by hormone receptor status are the small number of hormone negative types of breast cancer and that the stratified analysis by hormone receptor status was unplanned in the DBCPT. Logistic regression based on a small number of cases with many adjustments for potential confounders and secondary analysis may produce a result by chance. The association between GI and ER negative tumours needs to be confirmed in a larger sample with an a priori hypothesis related to hormone receptor status.

We studied dietary intake of nutrients and GI during adulthood and it may be important to consider the associations between early life exposures and breast cancer risk. The impact of early-life exposures on breast cancer risk was evident in the epidemiological studies of women who were exposed to high levels of radiation during childhood and adolescence from the atomic bomb explosions in Hiroshima and Nagasaki. Breast cancer rates were higher among women who were 10 to 19 years of age when the bombs exploded but the highest risk was seen in women who were younger than 10 years of age at that time (182). These results suggest that the breast is susceptible to exposures during this period of cell differentiation and growth. Observational studies of soy intake during childhood, adolescence and adulthood have consistently shown early exposure to high intakes of soy foods was associated with a decreased risk of breast cancer and risk may be further reduced by intake as an adult (183-185). Another example of the impact of early-life exposures on breast cancer risk comes from The Nurses’ Health Study. They demonstrated that body fatness during childhood and adolescence was inversely associated with breast cancer risk and that this association was independent of adult adiposity (186). There have only been two studies of diet GI during adolescence and their results are inconsistent (78;86). Although early-life environmental exposures are important, there is insufficient evidence to determine if introducing a low GI during childhood and adolescence would be protective against breast cancer later in life.

5.3.2 Dietary data and GI values

Our dietary data was based on multiple food records collected at baseline and annually throughout the trial follow up, and likely provided a better estimate of diet GI than other studies that relied on FFQs. Evidence from the WHI and EPIC suggest that studies that
used four or seven day food records to estimate fat intake were able to detect a significant association between dietary fat and breast cancer risk, whereas analysis that used a FFQ did not (71;100). A nested case control study of dietary fibre and colorectal cancer also found that dietary data from four to seven day food records demonstrated an inverse association between dietary fibre and risk of colorectal cancer, but when the same analysis was conducted using dietary data from FFQ, no statistically significant association was observed (187). I am not aware of any studies that compared the ability of food records and FFQ to detect associations between GI with cancer risk.

There has been one validation study of diet GI determined from a FFQ in 121 Swedish men (188). The FFQ was developed to measure dietary factors in the Swedish population. The subjects kept two sets of 1-week food records, about six months apart, and completed a 96-item FFQ on two occasions, at baseline and after the food records. The measurement of mean diet GI from the two instruments were moderately correlated after adjustment for energy, $r = 0.54$. The mean diet GI from FFQs explained 29 percent of the variation of mean diet GI from food records. This result is similar to other validation studies of FFQ and food records for measurements of macronutrients (79-81;85;114). When they divided the data for mean diet GI into quintiles, FFQs classified 70% of subjects in the same or adjacent quintile as the food records. Although, the FFQs performed relatively well in comparison to food records in this population, FFQs have some measurement error which may attenuate risk estimates associated with diet GI in a way seen with dietary fat and risk of breast cancer (68;71;100). Furthermore, the FFQ used in the validation study was designed to measure dietary factors in the Swedish population and may provide a better estimate of diet GI than the FFQs used in the observational studies of GI and breast cancer risk which were focused on dietary fat.

Our study is the only one to use multiple food records to describe usual diet over a long follow up period of about 10 years. We used 3 day food records for a baseline measurement and an average of 17 food record days to measure post-randomization diet. Overall, our results became stronger with data from more food records in the analysis. In the stratified analysis, we saw a stronger risk estimate for diet GI and ER negative breast cancer with increasing numbers of food records. In the case only analysis we could detect a statistically significant relationship between GI and hormone negative breast cancer when we
used the mean of post-randomization and mean of all available food records, but not when our analysis was restricted to baseline data. Other studies that found associations between nutrients and disease risk using food records and not FFQs, collected 4- to 7- day food records over six months or a year (71;99;100;187). Our results suggest that multiple food records collected over many years also provide useful dietary data.

The distribution of GI values in our study was similar to other studies of GI and breast cancer risk (literature review table 2.1). Our median GI of 57 was comparable to the median value reported by other studies of 53 to 71. Our range of GI values as measured by the mean of the lowest to the highest quintile was 54 to 61, was also similar to the other studies. Compared to our pilot study of 50 food records with one food record per subject (methods 3.9), the distribution of diet GIs from the nested case-control study was narrower with more observations per subject. The SD for diet GI values in the pilot study was 5.3 and the SD for this study was 3.1 to 3.2 at baseline and 2.2 to 2.5 at post-randomization. Using multiple food records decreased the variability in diet GI and reduced the probability of a type II error (false negative).

Our methodology for assigning GI values to items on food records was developed based on reports from five groups who had compiled GI databases (140;143;189-191). We were strongly influenced by the work of Olendzki et al who developed a GI database to be compatible with NDS for use with 24 hour dietary recalls (140). They described in great detail the determination GI values based on CHO composition, product information and food preparation. All GI databases are subject to a fundamental limitation related to the determination of GI values for specific foods. Blood glucose response measurements are known to have within-subject, between-subject and between laboratory variation. The FAO/WHO has recommended a number of procedures to minimize these types of variation and we only used GI values from studies that measured blood glucose response and determined GI values in a manner consistent with the suggested protocol.

We enhanced the reliability of our diet GI values by using highly trained staff and following quality control procedures recommended by the EPIC study in 2008 to reduce inter-rater variation (143). The EPIC study noted statistically significant inter-rater differences in the assignment of GI values to foods on their FFQs. Therefore, we assigned GI values based on a consensus of two registered dietitians. Other than the EPIC study in
France (81), no other study mentioned whether they used dietitians to process their dietary data or described the methodology involved in assigning GI values.

We were the only study to use the most recent version of the International Tables of GI values published in 2008. They listed GI values for about twice as many food items as the former version published in 2002, used by most other studies. Despite access to new GI data, only 29% of our GI values came from tested values on foods that were directly comparable or similar to the foods in our database. Therefore, like other studies, we estimated or calculated the majority of GI values. We tried to assess the accuracy of methodology by examining correlations between diet GI and specific types of CHO. Our diet GI was moderately correlated with total fibre, starch, lactose and fructose, at both times points, in cases and controls ($r = 0.24 – 0.36, p < 0.0001$). These associations were similar in direction and magnitude to other studies of nutritional correlates with GI (107;109).

The assignment of GI values was somewhat constrained by the codes and descriptors within NDS. NDS codes and descriptors could not accommodate distinctions in the type of potato or rice. The GI of various types of these popular, starchy foods varies widely and we were required to estimate a mean value for the NDS codes for boiled or baked potatoes and cooked rice, which may not have accurately represented the food consumed. However, we could determine when potatoes were eaten cold in potato salad and modified our automated system to accommodate that situation. Our system for shared food codes allowed us to consider brand names of cookies, crackers and cereals with unique GI values.

NDS derived much of its nutrient database from the USDA food composition data. The data for macronutrients, such as total CHO, protein and fat, is complete but there is some missing data for sugars and fibre. As a result, our results for mono- and disaccharides may be underestimated and affect their correlations with GI.

GI contributed new information, not captured by nutrients, to the analysis of diet and breast cancer risk. Although we could predict about 50% of variation in diet GI in a regression model with total fibre, starch, lactose, fructose and sucrose, GI was not fully explained by these nutrients. Energy, protein and fat were also added to the model, but they were not significant. Determination of diet GI reflects something other than macronutrient intake. It is a composite measure based on a variety of CHO containing foods in the diet and as such GI may reflect an eating pattern.
5.4 Conclusions

I have achieved the four objectives of this project. The first objective was to calculate diet GI from multiple food records. We designed and implemented an automated system to assign GI values to food items on food records. We determined the diet GI from 13,554 food records collected from 660 subjects in a nested case-control study in order to investigate the association of GI and breast cancer risk. The second objective was to determine if GI was associated with breast cancer risk overall and stratified by hormone receptor status. We did not observe an association of GI with overall breast cancer risk or ER positive breast cancer, but we detected a significant positive association of GI with ER negative breast cancer after adjustment for total sugars. These results are based on a secondary analysis with a small number of cases and need to be confirmed in a larger sample. The third objective was to examine whether the associations of starch and total sugars with breast cancer risk observed in the DBCPT were independent of diet GI. Diet GI did not explain the associations observed between type of CHO and risk of breast cancer by hormone receptor status. When diet GI was included in the analysis, starch remained negatively associated with risk of ER positive breast cancer and the positive association between total sugars and risk of ER negative breast cancer became stronger. The fourth objective was to determine if diet GI was associated with breast cancer phenotype in a case-only analysis. Diet GI was strongly positively associated with hormone receptor negative breast cancer but was not associated with tumor size or nodal status.

Breast tumors differ biologically and clinically by hormone receptor status. Progress has been made in our understanding and ability to prevent ER positive breast cancer. However, little is known about the etiology of ER negative breast cancer which has a poor prognosis and often occurs in premenopausal women. The findings from this study and research on dietary patterns suggest that diet, including GI, may play a role in the development of ER negative breast cancer. Hormone receptor status may be more important than menopausal status or BMI to our understanding of the relationship between dietary factors and the development of breast cancer.
5.5 Future Work

The association between diet GI and hormone negative breast cancer needs to be confirmed in a larger sample. We continue to follow the subjects from the trial but further follow up is unlikely to yield enough cases of hormone negative breast cancer to confirm our findings. Therefore, it will be necessary for larger studies, with more hormone negative breast cancer cases and ideally with food record data, to replicate our study. I am aware of two studies that could potentially fulfill this goal. First, the Women’s Health Initiative Observational Study has identified 430 cases of hormone negative breast cancer and has dietary data from FFQs and food records (192). The other study is The Multiethnic Cohort Study of Diet and Cancer. They recruited subjects from five ethnic groups, including African Americans who are at increased risk for ER negative breast cancer. The Multiethnic Cohort Study of Diet and Cancer has identified 491 cases of hormone negative breast cancer cases and collected dietary data from FFQs and 24 hour recalls (55).

Animal studies can test specific hypotheses and may be helpful in examining the relationship between GI and hormone negative breast cancer. Although there are some limitations in extrapolating results to humans, animal models of non-hormonally responsive mammary tumors are available and could be helpful in advancing our understanding of this relationship.

We could contribute to the understanding of the biological effects of the GI and its relationship with hormone negative breast cancer. We have non-fasting serum samples from the subjects in the trial collected at baseline and annually during the trial follow up that could be analyzed to examine biomarkers of GI, such as lipids and adiponectin, that may be involved in estrogen independent mechanisms related to the development of ER negative breast cancer. Although our sample size is limited, our findings could generate hypotheses for research initiatives that may ultimately reduce the burden of ER negative breast cancer.

We plan to use our food record database and GI database to explore relationships between GI and dietary patterns. Research on dietary patterns suggests that “healthy” diets may be protective against ER negative breast cancer. It would be worthwhile to estimate the diet GI of dietary patterns related to breast cancer risk in order to identify any indirect effects of diet GI on breast cancer development.
As GI was not associated with ER positive breast cancer, the inverse association between starch and ER positive breast cancer requires further study. We will look for diet patterns that may explain the protective effect of starch on this major sub-type of breast cancer. We can also measure biomarkers in the stored serum samples to further our understanding of biological mechanisms that may be relevant.
Chapter 6: References


27. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51


59. Hamajima N, Hirose K, Tajima K et al. Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. Br J Cancer 2002;87:1234-45.


