2017

Development Of An Estrogen-Related Dietary Pattern And Lifestyle Score To Examine Breast Cancer Risk In Postmenopausal Women

Mark Andrew Guinter
University of South Carolina

DEVELOPMENT OF AN ESTROGEN-RELATED DIETARY PATTERN AND LIFESTYLE SCORE TO EXAMINE BREAST CANCER RISK IN POSTMENOPAUSAL WOMEN

by

Mark Andrew Guinter

Bachelor of Science
University of Iowa, 2012

Master of Public Health
University of Louisville, 2014

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Epidemiology

The Norman J. Arnold School of Public Health

University of South Carolina

2017

Accepted by:

Susan E. Steck, Major Professor

Anwar T. Merchant, Committee Member

Alexander C. McLain, Committee Member

Dale P. Sandler, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ACKNOWLEDGEMENTS

I would like to thank my family and friends for their unwavering support throughout the duration of my education. Without their encouragement I would not have achieved my many successes, or overcome my failures, in pursuit of a rewarding and fulfilling career in public health.

I also would like to express thanks to my mentor and committee chair, Dr. Susan Steck. Your invaluable guidance and direction has helped me to grow as an independent researcher, as well as a professional. I cannot overstate how fortunate I am to have had such a strong advocate for me, providing me the freedom to follow my interests. In addition to Dr. Steck, I would like to thank my remaining committee members: Dr. Anwar Merchant, Dr. Alexander McLain, and Dr. Dale Sandler. Your attentiveness and genuine involvement in my work has been much appreciated. My dissertation work, our various other collaborative projects, and my scientific development have undoubtedly benefited from your expertise.

This dissertation has been supported by the University of South Carolina's Office of the Vice President of Research through the Support to Promote the Advancement of Research and Creativity Graduate Fellowship Program; and in part by a grant from the National Institutes of Health – General Medical Sciences (T32-GM081740).

ABSTRACT

Introduction: Studies examining the association between individual dietary components and breast cancer have been inconclusive. The use of dietary patterns is a holistic approach which may yield stronger associations. We sought to develop a dietary pattern based on an estrogen metabolite (EM) profile hypothesized to increase breast cancer risk (high unconjugated estradiol and low ratio of 2- to 16-hydroxylated EMs (2/16 ratio)). This estrogen-related dietary pattern (ERDP) was examined for associations with postmenopausal breast cancer in two study populations and was incorporated into an estrogen-related lifestyle score (ERLS) with other modifiable risk factors for breast cancer. **Methods:** EM and dietary data from 653 postmenopausal women from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) were used to develop the ERDP. Reduced rank regression modeling was applied to identify food groups which explained the largest variation in the two EMs. The resulting dietary pattern was then applied separately in 28,304 and 37,752 women from PLCO and the Sister Study (SS), respectively, to examine associations with breast cancer using Cox proportional hazards models. The ERDP was incorporated into the ERLS with alcohol consumption, body mass index, and physical activity among PLCO participants. Increasing scores of the ERLS represent a lower combined exposure to estrogen with a total range of scores from 0 to 6. **Results:** ERDP scores contained foods with positively weighted intakes (non-whole/refined grains, tomatoes, cruciferous vegetables, cheese, fish/shellfish high in ω-3 fatty acids, franks/luncheon meats) and foods with negatively

weighted intakes (nuts and seeds, other vegetables, fish/shellfish low in ω-3 fatty acids, yogurt, coffee). In PLCO, a 1-unit increase in the ERDP score was associated with a 9%, 13%, and 13% increase in total (HR: 1.09, 95%CI: 1.01-1.18), invasive (HR: 1.13; 95%CI: 1.04=1.04-1.24) and estrogen receptor-positive (HR: 1.13, 95%CI: 1.02- 1.24) breast cancer, respectively. No association was observed in SS. PLCO participants in the highest ERLS category had a 34% (HR: 0.66; 95%CI: 0.56-0.78) reduction in risk of total breast cancer compared to the lowest category. **Conclusions:** A dietary pattern correlated with a high-risk estrogen profile was positively associated with postmenopausal breast cancer within the cohort in which it was derived. Potential differences in other risk factors or dietary assessment tools may explain differences in associations seen between PLCO and SS. Adopting a lifestyle that has a lower combined exposure to estrogen is likely effective in reducing the risk of postmenopausal breast cancer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2/16 ratio ..... Ratio of 2- to 16-hydroxylated estrogen metabolites

16OHE-1 ..... 16-hydroxyestrone

2OHE-1 ..... 2-hydroxyestrone

4OHE-1 ..... 4-hydroxyestrone

ACS ..... American Cancer Society

AHEI ..... Alternate Healthy Eating Index

AICR ..... American Institute for Cancer Research

aMed ..... Alternate Mediterranean Diet

B~FIT ..... Breast and Bone Follow-up to the Fracture Intervention Trial

BMI ..... Body mass index

CDC ..... Centers for Disease Control and Prevention

CI ..... Confidence interval

CMSB ..... Columbia Missouri Serum Bank

CUP ..... Continuous Update Project

CV ..... Coefficient of variation

DAG ..... Directed acyclic graph

DASH ..... Dietary Approaches to Stop Hypertension

DGA ..... Dietary Guidelines for Americans

DHEAS ..... Dehydroepiandrosterone

DII™ ..... Dietary Inflammatory Index

DQI-R ..... Diet Quality Index – Revised

DQX ..... NCI's Diet Questionnaire

E2 ..... Estradiol

EHBCCG ..... Endogenous Hormones and Breast Cancer Collaborative Group

EIA ..... Enzyme immunoassay

EM ..... Estrogen metabolite

EPIC ..... European Prospective Investigation into Cancer and Nutrition

ER ..... Estrogen receptor

ERDP ..... Estrogen-related dietary pattern

ERLS ..... Estrogen-related lifestyle score

FFQ ..... Food frequency questionnaire

GC/MS ..... Gas chromatography-mass spectrometry

HER2 ..... Human epidermal growth factor 2

HLIS ..... Healthy lifestyle index score

HR ..... Hazard ratio

HRT ..... Hormone replacement therapy

I3C Indole-3-carbinol

IGF ..... Insulin-like growth factor

LC/MS-MS ..... Liquid chromatography-tandem mass spectrometry

MeD ..... Mediterranean Diet

MET ..... Metabolic equivalent of task

MPED ..... My Pyramid Equivalents Database

MUFA ..... Monounsaturated fatty acids

NCI ..... National Cancer Institute

NHS ..... Nurses' Health Study

NIEHS ..... National Institute of Environmental Health Sciences

OC ..... Oral contraceptive

OR ..... Odds ratio

PA ..... Physical activity

PLCO ..... Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

PUFA ..... Polyunsaturated fatty acids

PR ..... Progesterone receptor

SD ..... Standard deviation

SEER ..... Surveillance, Epidemiology, and End Results

SFA ..... Saturated fatty acids

SHBG ..... Sex hormone binding globulin

SMC ..... Swedish Mammography Cohort

SS ..... Sister Study

RFS ..... Recommended Food Score

RIA ..... Radioimmunoassay

RR ..... Relative risk

RRR ..... Reduced rank regression

TNF-α ..... Tumor necrosis factor alpha

USDA ..... United States Department of Agriculture

VIP ..... Variable importance in projection

WC ..... Waist circumference

WCRF ..... World Cancer Research Fund

WHI ..... Women's Health Initiative

WHR ..... Waist-to-hip ratio

CHAPTER 1

INTRODUCTION

1.1 Statement of the Problem

Breast cancer, the most commonly diagnosed cancer among women worldwide, is a disease of strong hormonal influence.[1] An attenuation in the production of ovarian hormones is characteristic of the onset of menopause, which also corresponds to a change in disease risk.[2] Postmenopausal women, the population in which the highest proportion of incident breast cancer cases occur, have significantly lower circulating levels of estrogen compared to premenopausal women.[2,3] Many of the well-established factors associated with breast cancer, such as lactation, age at menarche, and parity are significantly associated with estrogen metabolism.[4–6] Additionally, serum and urinary levels of estrogen metabolites (EM) have been shown to be consistently associated with postmenopausal breast cancer risk in prospective investigations.[4,7–12] Therefore, modifiable lifestyle risk factors for postmenopausal breast cancer that are associated with estrogen metabolism may present opportunities for primary prevention.

There are many nutrition-related lifestyle factors that have been identified with sufficient evidence that influence the development postmenopausal breast cancer.[2,13] Both sides of the energy balance equation, excess intake in the form of adiposity and greater energy expenditure in the form of physical activity (PA), show evidence of a positive and inverse association with postmenopausal breast cancer, respectively.[2,13] Consumption of alcohol has also been shown to increase breast cancer risk.[2,13] Using indices to assess

modifiable lifestyle factors as one aggregate score has been promising in identifying associations with breast cancer risk.[14–16] The study of dietary factors, however, with the exception of alcohol, has yielded conflicting results in relation to breast cancer risk.[2,13,17–23] Other individual dietary factors, such as non-starchy vegetables and foods containing carotenoids have limited but suggestive evidence of an association with breast cancer according to the latest report by the World Cancer Research Fund (WCRF) and the American Institute for Cancer Research (AICR).[24] Furthermore, most other food components (e.g., fiber, fruit, and total fat intake) have such limited or conflicting evidence, the report deems their association with breast cancer to be entirely inconclusive.[24]

It is likely that the practice of studying dietary components in isolation is contributing to the inconclusive findings for associations with many diseases, according to United States Department of Agriculture's (USDA) Dietary Guidelines for Americans (DGA).[25] Nutrients are consumed in combinations, and many of these nutrients interact with one another with regards to digestion and metabolism. Therefore, it is beneficial to study diet in its entirety, as it is consumed, using dietary pattern analyses when investigating a potential association with breast cancer.[26] Emerging evidence has supported an association between some dietary patterns and incident breast cancer risk.[17,18,27] Many of the diets that have indicated an inverse relationship with breast cancer are characterized by high intakes of fruits and vegetables, and diets with increased risk typically have higher intakes in fat and animal products.[17,21,28] Although these components show no or weak associations with breast cancer when studied in isolation, they may influence risk when consumed as a part of a whole diet.

In order to address the inconclusive findings in the literature on diet and breast cancer, it may be beneficial to consider the mechanistic pathway by which a potential association may occur. Nutritional status, namely malnutrition, can influence many hormonal processes in women, such as the development of breasts, and the onset of both menarche and menopause.[29,30] Therefore, diet likely has some role in altering estrogen metabolism and subsequently breast cancer risk, similar to adiposity and PA.[13] A relatively new approach to dietary pattern analyses, reduced rank regression (RRR), allows the use of disease biomarkers, such as EMs, to develop a dietary pattern and then investigate its association with disease endpoints.[31] Previously, Fung et al. developed a dietary pattern correlated with serum levels of estradiol and estrone sulfate using RRR, but the pattern subsequently was not associated with breast cancer among postmenopausal women in the Nurses' Health Study (NHS).[32] However, application of the same estrogen-correlated dietary pattern in a Swedish cohort identified a positive association with incident breast cancer.[27] The potential effect of a dietary pattern based on estrogen metabolism, in isolation and in combination with other nutritional lifestyle factors, needs to be studied further in an attempt to identify primary prevention methods for public health intervention.

1.2 Purpose and Objectives

We used data from postmenopausal women in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) to develop a dietary pattern based on food groups that are correlated with serum estrogen levels. The estrogen-related dietary pattern

(ERDP) was applied to examine associations with postmenopausal breast cancer in PLCO. In order to examine the ERDP in a separate population from which it was developed, associations with breast cancer were also examined in the Sister Study (SS). Finally, in PLCO, the ERDP was combined with other lifestyle factors to assess the impact of an estrogen-related lifestyle score (ERLS) on the development of breast cancer.

An initiative of the National Cancer Institute (NCI), PLCO is a large population-based randomized trial to investigate the effect of regular cancer screenings on cancer mortality in men and women aged 55-74.[33] Control arm participants continued standard of care screening practices, while participants in the intervention arm underwent more frequent screenings over a six-year period. Enrollment took place from 1993-2001, with follow-up collected until 2015. The SS, sponsored by the National Institute of Environmental Health Sciences (NIEHS), is a large cohort study designed to examine genetic and environmental risk factors of breast cancer.[34] More than 50,000 sisters of breast cancer patients aged 35-74 were enrolled from 2003-2009 with follow-up data collection occurring every few years. Using these cohorts, we hypothesized that diets high in whole grains and vegetables, particularly dark green vegetables, and low in animal products would be characteristic of low ERDP scores. We expected the ERDP scores to be associated with postmenopausal breast cancer independently, and as a part of the ERLS. Our study aims were as follows:

Aim I: To derive a dietary pattern based on estrogen metabolites and apply it to examine risk of postmenopausal breast cancer.

Fifteen EMs have been assayed using baseline serum samples from a nested case-control study of postmenopausal women enrolled in PLCO.[35] In Aim #1, we identified

food groups most strongly associated with EMs to create a dietary pattern that characterized a woman's diet based on its hypothesized cumulative estrogenic properties. Two EMs with sufficient evidence of an association with postmenopausal breast cancer were used in the development of the ERDP. Previous research on dietary patterns and breast cancer has been inconclusive, however, most of the patterns have not considered disease mechanisms specific to breast cancer.[17,21,28,36–40] Evidence from two studies that utilized an estrogen correlated dietary pattern have been mixed.[27,32] Using data from all postmenopausal women in PLCO's intervention arm, the ERDP was used to prospectively assess its association with overall postmenopausal breast cancer and by estrogen receptor (ER) subtype, with consideration of potential effect modifiers. We aimed to answer the following questions under Aim #1:

1. What food groups are most strongly correlated with serum EMs?
2. How much of the variation in serum EMs are explained by the ERDP?
3. Is there an association between the ERDP and overall breast cancer risk among postmenopausal women?
4. Does the association between the ERDP and postmenopausal breast cancer vary by ER subtype?
5. Is the association between the ERDP and postmenopausal breast cancer modified by other estrogen-related risk factors (e.g., obesity, parity, alcohol consumption, hormone replacement therapy (HRT))?

Aim II: To examine the association between the ERDP and postmenopausal breast cancer in an external study population from which it was developed.

A potential association between the ERDP derived in Aim #1 and breast cancer was investigated further in Aim #2, using prospective data from postmenopausal women enrolled in SS. Use of data from the SS allowed for the examination of the association between the ERDP and breast cancer in a different population from the one in which it was derived as a potential validation study for any observed associations in PLCO. Similar to Aim#1, the association was investigated for overall breast cancer and by ER subtype, with consideration of potential effect modifiers. The following questions pertained to Aim #2:

1. Is there an association between the ERDP and overall breast cancer risk among postmenopausal women?
2. Does the association between the ERDP and postmenopausal breast cancer vary by ER subtype?
3. Is the association between the ERDP and postmenopausal breast cancer modified by other estrogen-related risk factors?
4. Did the association between the ERDP and postmenopausal breast cancer differ between participants of SS and PLCO?

Aim III: To assess the relationship between an estrogen-related lifestyle score and postmenopausal breast cancer risk.

In Aim #3, a lifestyle score was developed using the ERDP in combination with other estrogen-related lifestyle factors known to be associated with postmenopausal breast cancer. Previous aggregate lifestyle scores have shown strong inverse associations

with breast cancer risk, but not always specific to postmenopausal breast cancer.[14–16] The scores have been based on cancer prevention recommendations from the WCRF/AICR,[14] or using investigator-defined components,[15] such as diet, physical activity, tobacco use, alcohol intake, and/or anthropometry.[16] Furthermore, to the best of our knowledge, no lifestyle scores have been developed to focus on a single disease mechanism, such as alteration of estrogen metabolism. Using data from all postmenopausal women in PLCO's intervention arm, the ERDP, body mass index (BMI), alcohol use, and PA were used to characterize an ERLS. The ERLS was investigated in relation to overall postmenopausal breast cancer and by ER subtype, with consideration of potential effect modifiers. We aimed to answer the following questions under Aim #3:

1. Is there an association between the ERLS and overall breast cancer risk among postmenopausal women?
2. Which components of the ERLS are the strongest contributors to a potential association with breast cancer?
3. Does the association between the ERLS and postmenopausal breast cancer vary by ER subtype?
4. Is the association between the ERLS and postmenopausal breast cancer modified by other estrogen-related risk factors (e.g., parity, HRT)?

1.3 Significance of the research

Previous research on the association between diet and breast cancer has been inconclusive. The research performed in this dissertation is innovative in that it addressed a disease mechanism specific to breast cancer by identifying a dietary pattern associated

with EMs. The association between the dietary pattern and breast cancer incidence was assessed using two large, federally-sponsored prospective cohort studies. The EM data in PLCO were generated using accurate and sensitive methods for assaying the low concentrations present in postmenopausal women, allowing for minor discrepancies in EM levels from dietary exposures to be quantified.[41] The previously derived estrogen-correlated dietary pattern has shown mixed but promising associations between diet and postmenopausal breast cancer.[27,32] The methods employed in this dissertation are believed to have improved on the previous study by creating a newly derived pattern using different EMs, which may be more representative of breast cancer risk than the previously-used parent estrogens. Furthermore, a more sensitive assay was used in measurement of the EMs that may be particularly meaningful considering the low levels of EMs present in postmenopausal women. Evaluation of the ERDP in multiple study populations and as a part of the ERLS attempted to quantify the magnitude of the effect of estrogen-related nutritional factors on breast cancer in postmenopausal breast cancer.

1.3.1 Public health impact

In recent years, advances in the treatment of breast cancer have led to a substantial reduction in mortality rates.[42] However, 1 out of every 8 women born in the U.S. will be diagnosed with breast cancer in their lifetime.[2] As increasing worldwide industrialization and urbanization has resulted in rising global incidence rates, the need for primary prevention methods for breast cancer is of upmost importance.[24] The collaborative 2012 Breast Cancer Campaign, made up of over 100 international experts in breast cancer, identified the need for sustainable lifestyle prevention methods as one of the 10 most important gaps in translational breast cancer research.[43] The results from this dissertation

contribute to the literature on dietary habits, alone and in combination with other lifestyle factors, with the intent to lower future breast cancer incidence among postmenopausal women.

1.3.2 Role of diet in breast cancer is inconclusive

Results from research examining dietary exposures and breast cancer risk have been inconsistent, although a modest effect has been suggested.[2,13,22] According to the WCRF/AICR, the only nutritional factors with conclusive or probable evidence of an association with postmenopausal breast cancer are alcohol consumption, body and abdominal fatness, and PA.[13,24] There is suggestive evidence of an effect from starchy vegetables, foods containing carotenoids, and diets high calcium; however, the evidence and biologic plausibility are lacking.[24] Studying associations between diet and breast cancer may be inconclusive due to the heterogeneity in disease characteristics for pre- and post-menopausal women and hormonal subtypes.[22,44]

However, it has been suggested that dietary habits and other lifestyle behaviors are often adopted together, and may have a collective effect on cancer risk.[45,46] There is evidence that choosing to eat healthy foods together, thus improving overall diet quality, is associated with reduced cancer risk, such as with the Mediterranean diet (MeD).[17,40,47,48] When higher dietary quality is measured by patterns based on associations with certain markers of disease risk, such as inflammation or estrogen metabolism, associations with breast cancer have been identified in some studies,[27,49,50] but not all.[32,51] Together, the research indicates that when diet is evaluated as the sum of individual components, which likely interact with each other, a dietary influence on the development of breast cancer in postmenopausal women is more likely to be found than

when examining individual dietary factors. Individual components may influence disease risk in multiple different mechanistic pathways or by predominantly converging on a single pathogenic pathway.

1.4 Study outline

The overall goal of this dissertation was to investigate the relationship between diet and postmenopausal breast cancer, in addition to other lifestyle factors, with consideration of estrogen metabolism as a possible disease mechanism. The rationale and significance of the research proposed is outlined in Chapter 1. Chapter 2 details the background for all of the potential relationships considered, as well as important confounders. The current knowledge on the associations between estrogen and breast cancer, diet and estrogen, and diet and breast cancer are presented, along with other risk factors for breast cancer, in Chapter 2. The background review served as the rationale for the methods proposed in Chapter 3. Descriptions of the study populations and the analytic approaches used can be found in Chapter 3. The results from each of the three dissertation aims are reported in Chapters 4, 5, and 6 separately. Those chapters are written in a manner so that each one represents a publishable manuscript. Chapter 7 is a synthesis of the three aims and discussion of the collective results.

CHAPTER 2

BACKGROUND

2.1 Estrogen and breast cancer

There are a number of well-established risk factors for breast cancer that, when looked at collectively, highlight the presence of a hormonal influence on the development and prognosis of breast cancer.[2,4,12,13] Menopausal status, age at menarche, parity, age at menopause, adiposity, and alcohol intake, to name a few, all have a commonality other than their association with breast cancer risk.[4,52] These factors, described in more detail in section 2.4, have been shown to have a significant relationship with endogenous estrogen levels.[4,12] For example, adipose tissue is recognized to have endocrine functionality and has been shown to promote the synthesis of estrogens via high expression of aromatase, especially in postmenopausal women.[53–56] Estrogen itself, measured either in serum or urine, has repeatedly shown a positive association with postmenopausal breast cancer risk, as discussed below (section 2.1.3).[11,35,57] Although estrogen is the strongest sex hormone correlate of breast cancer, there is evidence of an association with breast cancer for various other sex hormones, including androstenedione, dehydroepiandrosterone (DHEA) and testosterone.[58–61]

The role of hormones in breast cancer risk extends beyond steroidal hormones. In addition to estrogen, adipose tissue has the capability of producing non-steroidal hormones, such as leptin and adiponectin, which can influence mammary carcinogenesis both directly and indirectly.[62,63] Leptin has been shown to have a proliferative effect on

breast cancer cells through enhancement of multiple signaling pathways, whereas adiponectin has been suggested to down-regulate cell proliferation and even induce apoptosis.[62,64] In addition, there are other breast cancer risk factors that are associated with important hormones other than estrogen in the development of cancer, such as prolactin in breastfeeding.[13,65]

Perhaps the most compelling argument to characterize breast cancer as a disease of major hormonal influence is the recognition of four distinct molecular subtypes of breast cancers.[66] Characterized by the presence of hormonal receptors for estrogen and progesterone (PR), in addition to levels of human epidermal growth factor 2 receptors (HER2), each molecular subtype differs with regard to incidence rates, risk profiles, and prognosis.[2,67] The association between estrogen and many breast cancer risk factors, along with its influence on the way the disease manifests and progresses, highlight the importance of incorporating the extensive influence of estrogen in investigations of breast cancer.

2.1.2 Laboratory methodology

The majority of epidemiologic research on estrogen metabolism in relation to breast cancer risk over the last 20 years was conducted using a radioimmunoassay (RIA) or an enzyme immunoassay (EIA).[11,57,68] The RIA method is known as an extraction assay because it includes an extraction and subsequent purification step, but also requires a large volume of the serum sample, thus limiting its application.[69] To help test the association with estrogen and breast cancer in large epidemiologic studies, EIA was developed because of its rapid and inexpensive application.[11] The assay can be applied to both urine and serum samples and was called a “direct” assay because it did not involve

any purification or extraction steps in the process.[68] Although useful in ranking individuals, EIA was inadequate for absolute measurements of hormones from samples.[69] A study evaluating EIA was able to show the method was reproducible in premenopausal women, with a coefficient of variation (CV) between 8-14% for urine samples.[68] However, when EIA was applied to urine samples from postmenopausal women the mean levels increased over 50% from the 4-month interval to the 12-month interval of reproducibility.[68] Comparing against the standard at the time of publication, gas chromatography-mass spectrometry (GC/MS), EIA was shown to have a lower specificity and reproducibility.[41,68] Quantitative comparison studies had shown that although they were sensitive among premenopausal women, RIA and EIA had poor specificity and accuracy, likely a result of cross-reactivity and batch-to-batch variation of the antibodies in the urine samples.[70] Cumulative evidence clearly showed that more precise and accurate assay methods were needed to assess the relationship between estrogen metabolism and breast cancer, particularly among postmenopausal women.[11,41,70,71]

The use of GC/MS in large scale studies is impractical because of its cost and arduous application, however mass spectrometry assays have been shown to be most accurate and reproducible.[70] Coupling the need for an inexpensive, accurate and reproducible method with the increasing recognition of the influence of estrogen metabolism in all its forms and pathways, liquid chromatography-tandem mass spectrometry (LC/MS-MS) was developed for urine and serum samples.[41,70] Comparing urine samples from 430 women using EIA and LC/MS-MS, absolute concentrations of 2-hydroxyestrone (2OHE-1) and 16-hydroxyestrone (16OHE-1) were consistently higher in

EIA.[72] The difference of the assays by menopausal status was particularly striking, with mean concentrations for premenopausal 2-4 times higher and for postmenopausal 7-12 times higher when comparing EIA to LC/MS-MS.[72]

Using LC/MS-MS, researchers can concurrently measure 15 EMs in an accurate and reproducible method with enough sensitivity to detect the low levels present in postmenopausal women.[11] In the nested case-control of postmenopausal women enrolled in PLCO used in the present proposal, blind quality control serum samples were shown to have a CV <5% for all 15 EMs using LC/MS-MS.[35] Furthermore, the CV for the parent estrogens, estradiol and estrone, were <3% in the samples.[35] In the previous study of an estrogen-correlated dietary pattern by Fung et al., the RIA method was utilized and only estradiol and estrone sulfate were assayed with reported CVs<15%.[32] The LC/MS-MS method has been shown to have an intraclass correlation greater than 95% among postmenopausal women and the lowest limit of detection with reliable and reproducible estimates is between 1-2 pmol/L from serum samples.[35,73,74] For reference, the measured levels of estradiol in postmenopausal women who are not currently undergoing HRT typically range from 0-117 pmol/L.[69,75,76] The current and previously referenced evidence supports the use of LC/MS-MS as an accurate, sensitive, and reproducible method to measure EM in postmenopausal women.

2.1.3 Evidence from observational studies

The precursors and downstream metabolites of estrogen have different physiologic effects as a result of their chemical structures.[4] Both parent estrogens, estradiol and estrone, are derived from the sex hormone, androstenedione (**Figure 3.1**). Androstenedione can be directly aromatized in estrone, but requires an additional step to

synthesize estradiol.[77] Androstenedione must first be reduced to testosterone, which can be subsequently aromatized to produce estradiol.[77] Estrone can be converted to estradiol, the most biologically active estrogen, by the 17β-hydroxysteroid dehydrogenase enzyme.[78] Once the parent estrogens have been synthesized, they may be metabolized down one of three, competing and irreversible pathways.[4] The three pathways are characterized by the carbon position (2, 4, or 16) that is hydroxylated by the cytochrome P540 enzyme.[4] The result of the hydroxylation produces catechol estrogens, which may undergo methylation to be further metabolized into methoxyestrogens.[4]

Early epidemiologic studies established a relationship between high levels of circulating estradiol and estrone with breast cancer in patients using case-control study designs.[57] However, due to the potential for reverse causality, it was unclear whether the higher levels among cases were markers of disease risk or of the presence of disease. The estrogen hypothesis was studied further in large scale prospective studies, starting around 1990.[57] An international group called the Endogenous Hormones and Breast Cancer Collaborative Group (EHBCCG) conducted a meta-analysis of circulating hormones from nine prospective studies of postmenopausal women not using exogenous hormones, including 663 cases and 1765 controls.[57] Results showed significant associations with breast cancer comparing the highest quintile to the lowest for all hormones (estradiol, free estradiol, non-sex hormone binding globulin (SHBG) bound estradiol, estrone, estrone sulfate, androstenedione, DHEA, DHEA sulphate, testosterone).[57] Most effect estimates remained significant even after adjustment for estradiol, which was correlated with all hormones investigated.[57] The highest effect estimates were for free estradiol (relative risk (RR): 2.58; 95% confidence interval (CI): 1.76-3.78) and non-SHBG (RR): 2.39, 95%

CI: 1.62-3.54).[57] Apart from SHBG (RR: 0.66; 95% CI: 0.43-1.00), all associations were in the positive direction.[57] The inverse association between SHBG and breast cancer risk is hypothesized to be a result of its role in reducing circulating bioavailable estradiol.[8] Since then, EHBCCG conducted an updated meta-analysis comparing results of eighteen different prospective studies with consideration of the assay method used, while excluding women currently taking exogenous hormones.[69] The hormones of interest were estradiol, estrone, and testosterone. All 3 hormones, across all three assay methods (extraction, direct, and mass spectrometry), were significantly associated with postmenopausal breast cancer risk, with the exception of testosterone measured by mass spectrometry.[69] Again comparing the highest quintile to the lowest, the effect estimates ranged from 1.46 to 2.46.[69] Combining results from all assay methods, individuals in the highest quintiles experienced around twice the risk compared to the lowest quintile for estradiol (RR: 2.15; 95% CI: 1.87-2.46), estrone (RR: 1.81; 95% CI: 1.56-2.10), and testosterone (RR: 2.04; 95% CI: 1.76-2.37).[69] Since the most recent EHBCCG meta-analysis has been published, results from three prospective studies of postmenopausal women using LC/MS-MS have corroborated their results for the parent estrogens (estrone and estradiol), with unconjugated estradiol consistently showing the largest magnitude of an effect on breast cancer risk.[35,79,80] Data from the nested PLCO study to be used in the present proposal showed a doubling of risk comparing the highest and lowest decile for unconjugated estradiol (hazard ratio (HR): 2.07; 95% CI: 1.19-3.62) using serum samples.[35]

While there is an established relationship between circulating parent estrogens and breast cancer risk among postmenopausal women, there is not as much research on

other estrogen metabolites, partly due to limitations of previous laboratory assay methods.[11,57] As previously stated, the parent estrogens may be hydroxylated down one of three different metabolic pathways.[11] It has been hypothesized that shifts in these competing pathways may influence breast cancer risk.[11] Initial research in a case-control study had shown that breast cancer patients had 60% higher circulating levels of 16OHE-1 than controls, whereas 2OHE-1 and 4-hydroxyestrone (4OHE-1) levels were similar across the two groups.[11] Further research concluded the 2OHE-1 and 16OHE-1 were competing metabolic pathways representative of breast cancer risk, with women having a higher ratio of 2OHE-1 to 16OHE-1 (2/16 ratio) experiencing reduced risk of breast cancer.[11] Results of studies investigating the 2/16 ratio using EIA were inconsistent.[11]

However, in studies using the advanced LC/MS-MS to measure estrogen metabolism in postmenopausal women, results have more consistently shown a reduction in risk with increasing 2/16 ratio when looking at all estrogen metabolites combined, not only estrone and its hydroxylated forms.[35,79,80] In a nested case-control study from the Columbia Missouri Serum Bank (CMSB), comparison of 215 postmenopausal cases and 215 matched controls yielded a non-significant reduction in risk for the 2/16 ratio (odds ratio (OR): 0.63; 95% CI: 0.35-1.12) comparing the highest to the lowest quintiles.[80] Data from a larger case-cohort from the Breast and Bone Follow-up to the Fracture Intervention Trial (B~FIT), including 407 postmenopausal cases and 487 controls, identified a significant difference in risk comparing the highest and lowest quintile for the total 2/16 ratio from serum samples.[79] Women in the highest quintile had a 40% reduction in risk of breast cancer (HR: 0.60; 95% CI: 0.40–0.90).[79] In PLCO's nested case-control study, a similar reduction in risk was observed across the interdecile range of

the 2/16 ratio before adjustment for unconjugated estradiol (HR: 0.62; 95% CI: 0.45-0.86) and retained a similar magnitude of association, although insignificant, after adjustment (HR: 0.69; 95% CI: 0.47-1.02).[35]

The ratios of other EMs have also been investigated after the advent of LC/MS-MS. While results have been inconsistent with regards to statistical significance, the direction of the effects has been consistent throughout all three LC/MS-MS studies.[11] In addition to the 2/16 ratio, there is evidence the 2/parent estrogen ratio is associated with reduced risk in postmenopausal women.[35,79,80] In the CMSB study, the 2/parent ratio yielded a non-significant reduction in risk (OR: 0.63; 95% CI: 0.35, 1.12).[80] Similarly, results from B~FIT showed a non-significant reduction in risk of a similar magnitude (HR: 0.69; 95% CI: 0.46-1.05), but the test for trend was statistically significant (p=0.01).[79] In the PLCO population to be used in the present analysis, the 2/parent ratio was associated with reduced risk before (HR: 0.66; 95% CI: 0.51-0.87) and after (HR: 0.72; 95% CI: 0.52-1.00) adjustment for unconjugated estradiol.[35] In fact, data from PLCO yielded significant effect estimates for both the 2/16 and 2/parent ratios, but not unconjugated estradiol, when all three were entered into the model at once.[35] In the same PLCO study, the ratio of the 4-catechols to the 4-methylated catechols (4/4-methylated) was positively associated with postmenopausal breast cancer.[35] This supports laboratory evidence indicating the instability of 4-cathechol DNA adducts can be blocked by methylation.[81] However, other observational evidence has failed to support the findings from PLCO.[11]

2.1.4 Potential mechanisms

Collectively, the results from epidemiologic studies suggest increased circulating parent estrogens, particularly estradiol, is associated with an increase in postmenopausal breast cancer risk. Furthermore, it appears that enhancement of the 2-hydroxylation pathway, compared to both the 16-pathway and parent estrogens, is characteristic of a reduction in the risk of postmenopausal breast cancer. The mechanisms behind the influence of estrogen on breast cancer risk are not completely understood, and may act both independently and dependently through their receptors.[8] There is evidence of carcinogenic effects of estrogen in mammary tissue through multiple pathways from animal and human studies.[4,8,82,83] Treatment of mice with estrogen has been shown to be positively associated with mammary tumors.[82] In mature human breast tissue, there is evidence estrogen increases the rate of cellular proliferation.[8,83] In vivo and in vitro studies have shown downstream that metabolites of estrogen can lead to unstable adducts of adenine and guanine in DNA, consequently leading to mutations.[82] Conversely, other quinones produced in estrogen metabolism can establish a redox cycle, resulting in reactive oxygen species that can have detrimental oxidative effects on DNA.[8]

A pathway-specific investigation of estrogen metabolism may help to elucidate how the metabolite ratios can potentially affect breast cancer risk. The increase in 2/parent and 2/16 ratios are indicative of a possible protective effect of metabolites in the 2-hydroxylation pathway.[4,8,82] The downstream 2-hydroxylated metabolites have been shown to have a lower affinity for estrogen receptors, possibly due to a decreased hormonal effectiveness compared to estradiol.[4] There is some evidence that metabolites in the 2-hydroxylation pathway inhibit cellular growth and proliferation and are

associated with apoptosis.[4] On the contrary, metabolites from the 16-hydroxylation pathway have been shown to exhibit carcinogenic and genotoxic properties.[84] Mouse models with treatment of 16OHE-1 have resulted in spontaneous DNA synthesis in mammary epithelial cells.[4] Additionally, cancerous mammary tissue has been reported to have nearly eight times the amount of 16OHE-1 compared to fat tissue in the breast.[4] The competing nature of the 2- and 16-hydroxylation pathways, and their relative cellular effects, can help to explain why higher 2/16 and 2/parent ratios are associated with a reduction in postmenopausal breast cancer risk.

2.2 Diet and estrogen

In order to reduce the incidence of postmenopausal breast cancer, it is imperative to identify primary prevention methods, such as dietary intervention targets. While circulating levels of estrogen are an established risk factor, and many other risk factors are associated with estrogen metabolism, the evidence of a link between diet and estrogen is scarce.[2,13] The hypothesized relationship between diet and estrogen first originated in an attempt to explain results from ecological and migrant studies. Women from Eastern regions of the world experience much lower rates of breast cancer than Western women, possibly due to vast differences in diet.[85] When women migrate to the U.S., their disease risk begins to parallel that of U.S. born women.[86,87] A comparison of White U.S. women with Asian immigrants reported a 3-fold increase in plasma estradiol and lower fecal excretion of estrogen among Whites, hypothesized to be reflective of differences in diet.[88] Supporting literature on diet and estrogen metabolism is scarce, highlighting the need for

more investigations into the effects of single dietary components as well as dietary patterns on EMs.

2.2.1 Single dietary components and estrogen

One of the earliest published studies of diet and estrogen investigated differences in plasma, fecal, and urinary excretions between vegetarians and omnivorous women.[89] Over 4 months of follow-up, fecal excretion of estrogen was higher among vegetarians ($p<0.001$), and plasma estrone and estradiol levels were negatively correlated with fecal excretion of estrogen ($p=0.005$).[89] The researchers concluded that a vegetarian diet that led to larger excretion of estrogen, and subsequently lower plasma levels of estrogen, may be reflective of low intakes of fat and high fiber. Investigations of animal products in relation to estrogen metabolism support a positive relationship between fat intake and estrogen. A cross-sectional study of 766 postmenopausal women reported 13% lower mean plasma levels of SHBG in women in the highest quartile of red meat consumption compared to the lowest, with a significant test for trend ($p<0.01$).[90] Women in the highest quartile of dairy product consumption from the same study had 15% and 14% higher levels of total and free estradiol compared to the first quartile, again with significant trends ($p=0.02$ and $p=0.03$, respectively).[90] The observed association may be a result of the hormones that are present in the milk consumed, however, it has been suggested that the levels in milk are too low to have a physiological effect and may become inactive following digestion.[91,92]

Regarding dietary fiber, a mostly consistent inverse association with circulating estrogen has been shown in premenopausal women.[93–98] In postmenopausal women, however, the evidence of an association is not as strong. A study of 291 women in a

dietary intervention trial reported reduced serum bioavailable estradiol ($p<0.01$) and total estradiol ($p <0.05$) concentrations as a result of increased fiber intake after one year of follow-up.[99] Using data from Hispanic women in the Multiethnic Cohort study (MEC), differences of -22% ($p=0.023$) and -17% ($p=0.045$) for serum estrone and estradiol, respectively, were observed when comparing postmenopausal women in the highest quintile of fiber intake to the lowest.[100] It has been hypothesized that steroid hormones bind to certain types of fiber, which could also explain the increased fecal excretion of estrogen among vegetarians.[89,101] However, results from larger observational studies, mostly of cross-sectional design, have reported no association between dietary fiber and estrogen or other hormones.[102–106]

One of the more frequently studied dietary components in relation to estrogen is fat intake because of the established relationship between adipose tissue and estrogen synthesis. Results from intervention studies in postmenopausal women reported a significant reduction of plasma estradiol after lowering dietary fat,[107] a reduction in estradiol after a low-fat, high carbohydrate diet,[108] and lowered urinary excretion of estrone after participation in a low-fat intervention with high ω-3 fatty acid intake.[109] In a subset of postmenopausal women with plasma samples from the Women's Health Initiative (WHI) Dietary Modification Trial, which was designed to assess the relationship between a low-fat diet and breast cancer risk, a reduction of estradiol (relative change in geometric mean: 0.85; 95% CI: 0.72-1.00) and increase in SHBG (relative change: 1.09; 95% CI: 1.03-1.16) was observed among the intervention group.[110] A meta-analysis of 13 low-fat intervention studies reported a pooled estimate of

a 23% reduction (95% CI: 18.1%-27.7%) in circulating estradiol post baseline among postmenopausal women in the intervention groups.[111]

Results of observational studies examining associations between dietary fat and estrogen have been less conclusive. Cross-sectional studies have repeatedly shown no association between dietary fat intake and hormonal concentrations in postmenopausal women.[103,104,112,113] It is possible that weight loss mediated the association between a decrease in dietary fat and circulating estrogen observed in intervention studies. However, a cross-sectional analysis from the NHS reported 4.3% lower plasma estradiol (95% CI: 0.2%-8.3%) for every 5% decrease in energy from fat intake among 384 postmenopausal women.[102] In a Japanese study of postmenopausal women, baseline serum estrone was positively associated with baseline percentage of calories from total fat intake (p=0.04), and borderline significantly associated with monounsaturated fat (MUFA) and polyunsaturated fat (PUFA) (p=0.05 for both) intakes.[114] The same study reported significant positive associations between DHEAS from serum samples with percentage of energy from total fat (p=0.007), saturated fat (SFA) (p=0.009), MUFA (p=0.006), and PUFA (p=0.04).[105,114] Although a relationship between dietary fat and estrogen is still inconclusive, it is possible an observed association is a combination of effects from weight loss, reduced animal product intake, and increased fiber from plants.

In addition to vegetarianism or increased fiber from plants, other plant products have been associated with estrogen metabolism. Indole-3-carbinol (I3C), abundantly found in cruciferous vegetables, has displayed anti-estrogenic properties.[115,116] Consumption of I3C extracts in 10 women for 2 months resulted in a 0.26 $nmol/mmol_{creatine}$ decrease (95% CI: 0.06 to 0.46) in urinary estradiol after

intervention.[117] Decreases were also seen in estrone, estriol, and 16OHE-1 along with an increase in 2OHE-1, indicative of a beneficial alteration in estrogen metabolism.[117] An intervention study of cruciferous vegetables, particularly broccoli, found a 0.08 increase (95% CI, 0.02–0.15) in the 2/16 ratio from urine samples for each 10-g/day increase in cruciferous vegetables, showing a beneficial shift in estrogen metabolism.[118] In another intervention study of 13 premenopausal women, consumption of a powder from dried cruciferous vegetables increased the mean 2/16 ratio from 1.25 to 2.28 (p=0.01) using urine samples.[119] This association with serum or urinary estrogens has failed to be replicated in epidemiologic studies, likely due to the low consumption levels of cruciferous vegetables in some populations.[120] However, in studies of tumor cells there is evidence that cruciferous vegetables can shift estrogen metabolism in a favorable manner, particularly in reference to the 2-hydroxylation pathway.[121] The shift towards the 2-hydroxylation pathway is possibly a change in the relative production of cytochrome P540 proteins, resulting from exposure to I3C, which influences the metabolic pathways of the parent estrogens.[8,122]

As a result of the drastic differences in breast cancer rates among Asian countries and the U.S., and the relative differences in diet, soy intake has been hypothesized to beneficially affect estrogen metabolism.[88,123,124] Isoflavones, a type of phytoestrogen contained in soy, may alter endogenous estrogen metabolism and have the ability to bind to estrogen receptors.[125] In experimental studies of premenopausal women, those consuming increased soy products had decreased urinary estradiol, estrone, 16OHE-1, and a significant increase in 2OHE-1.[126,127] However, an intervention study of 97 postmenopausal women reported no change among urinary SHBG or estradiol after

consuming a high-soy diet for four weeks.[128] Collective evidence from a meta-analysis of intervention studies concluded no statistically significant effects of soy or isoflavone consumption on levels of estrone or SHBG among postmenopausal women, with similar results among premenopausal women.[129] A modest, non-significant increase in estradiol in the soy consumption groups (14%, $p = 0.07$) was reported. It is possible the null results were due to a failure to take into account the assay used to measure estrogen in the pooled analysis, and the authors used a funnel plot to show studies finding extreme increases or decreases in estradiol may have been excluded from their analysis.[129]

Data from observational studies regarding soy intake and estrogen metabolism are limited. A cross-sectional study of Asian-American women reported significant 16% higher urinary levels of 2OHE-1 ($p_{trend}=0.02$) accompanied by 11% lower levels of 16OHE-1 ($p_{trend}<0.01$) comparing the highest versus lowest tertiles of soy consumption.[130] Results from another study among postmenopausal Chinese women showed 15% lower plasma levels of estrone among the highest quartile of soy consumers compared to the lowest.[112] In a study of predominantly White, British women, no association was found between plasma EMs and soy milk intake for pre- and post-menopausal women.[131] There is evidence of a reduction in estradiol which depends on the presence of certain polymorphisms, suggesting a gene-diet interaction which could help to explain the differences observed by race/ethnicity.[132] It has also been hypothesized the large discrepancies in intake and lifetime exposure to soy explain the differences observed between Asian and non-Asian populations.[124]

Alcohol intake, another dietary factor that could explain the large differences in breast cancer incidence rates across the world, has strong evidence of an association with

estrogen metabolism.[2,12,133,134] A positive association between alcohol intake and circulating levels of hormones has been demonstrated in premenopausal women.[135,136] In a six-month cross-over trial of 34 premenopausal women, 30 g/day ethanol intake was associated with increased levels of urinary estrone by 15.2% ($p=0.05$), estradiol by 21.6% ($p=0.02$), and estriol by 29.1% ($p=0.03$).[137] In a prospective study of 66 premenopausal women, a modest but significant positive association was observed using Spearman correlation coefficients ($r=0.29$; $p<0.05$) between alcohol intake and serum estradiol concentrations.[138] There has been some evidence of a stronger effect among women using oral contraceptives (OC).[139] Studies of postmenopausal women have been more inconsistent.[135,140–142] In a randomized, controlled 6 week cross-over trial in which 40 and 30 g of alcohol consumption per day for men and women, respectively, for three weeks was compared to a three week abstinent period, plasma DHEAS increased but estradiol was not affected by alcohol intake, among 10 postmenopausal women.[143] In the reanalysis of 13 studies of postmenopausal women, all hormones measured were positively associated with at least 20 g/day of ethanol, with the highest difference in mean concentrations observed for DHEA sulphate (25%; $p<0.001$) compared to women who abstained from drinking.[12] Using data from nearly 2000 women enrolled in the European Prospective Investigation into Cancer and Nutrition (EPIC), pre- and postmenopausal women who consumed at least 25 g/day of ethanol had nearly 40% ($p<0.001$) and 20% ($p<0.001$) higher serum concentrations of estrone, respectively, compared to non-consumers.[136] Similar to a potential effect modification by OC in premenopausal women, there is evidence of a stronger association between alcohol intake and breast cancer among HRT users.[135]

2.2.2 Dietary patterns and estrogen

The literature on dietary patterns and estrogen metabolism are scarce, but there is some evidence of an association. Dietary data from postmenopausal women enrolled in the NHS showed associations with sex hormones in a cross sectional analysis of dietary patterns.[144] The Alternate Healthy Eating Index (AHEI), an *a priori* dietary pattern based on the USDA's DGA, was inversely associated with plasma estradiol ($p<0.001$) and positively associated with SHBG ($p=0.01$).[144] The results indicate a beneficial effect of better diet quality on estrogen metabolism, although results were attenuated after adjustment for BMI ($p=0.08$ and $p=0.37$, respectively).[144] Using principal component analysis to derive *a posteriori* patterns, the prudent pattern, characteristic of intake of plant products and whole grains, was not associated with EM.[144] The Western pattern, comprised of processed foods and animal products, was inversely associated with SHBG ($p=0.008$) before adjustment for BMI, but not after adjustment.[144] The Western pattern also was positively associated with total ($p=0.01$) and free ($p=0.006$) estradiol, but after adjusting for BMI, only the association with free estradiol remained statistically significant ($p=0.03$) when comparing the highest and lowest quintiles of the dietary pattern score.[144] The association for the Western pattern was replicated in a case-control of Mexican women, with authors reporting a 16.2% increase in the serum concentrations of free estradiol ($\beta=0.15$; 95% CI: 0.01-0.29) for every 1-unit increase in the dietary pattern score.[145] Although premenopausal women may be less sensitive to dietary estrogenic effects due to their higher mean circulating estrogen concentrations, an NHS investigation observed associations for the AHEI.[146] Women in the highest quartile of the AHEI had lower mean plasma levels of total estradiol (-6.7 %; 95% CI: -14.3% -1.5%;

p_{trend}=0.04) and androstenedione (-7.8%; 95% CI: -15.4%-0.4%; p_{trend}=0.03) compared to the first quartile, although no associations were evident for adherence to the Dietary Approaches to Stop Hypertension (DASH) or the alternate Mediterranean Diet (aMeD).[146] The (MeD) and its alternate form (aMeD) are based on the dietary characteristic of people living in that region, as opposed to dietary guidelines like the previously mentioned *a priori* indices. The MeD is usually high in fruits and vegetables, legumes, oils, and other foods that result in a higher proportion of MUFA and PUFA compared to saturated fats.[47] While the previously mentioned study reported no association for the aMeD, an intervention study using the MeD reported a roughly 40% decrease in total urinary estrogen levels ($p<0.02$) in postmenopausal women, showing some anti-estrogenic properties.[147] Collectively, the published results show some evidence of associations between dietary patterns and estrogen metabolism, although results have been inconsistent.

2.3 Diet and breast cancer

An important lifestyle contributor to disease is diet, which has been estimated to be the second most preventable cause of cancer.[148] Prior research has indicated that 32% of all cancers may be avoided through proper dietary modification, with at least 1 in 5 cancer deaths preventable through diet.[149] However, cancers of differing anatomical sites are different diseases, as is their etiology. Information on lifestyle prevention measures, including diet, has been identified as one of the ten most important gaps in translational breast cancer research.[43] Although much of the research into diet and breast cancer has been inconclusive, it may be due to the heterogeneity of cancer subtypes, or due to the

relatively small effects from single dietary components, that may be magnified when studying diet holistically.

2.3.1 Single dietary components and breast cancer

According to the most recent 2017 Continuous Update Project (CUP) of the WCRF/AICR's Second Expert Report, alcohol intake is the only dietary factor designated to have a "convincing" association with an increased risk of breast cancer.[24] The report cited a recent meta-analysis of 22 prospective cohort studies, identifying an 9% increase in postmenopausal breast cancer for every 10g of ethanol consumed each day.[24] In an additional pooled analysis of including over 33,000 incident breast cancer cases, a significant increase in risk of 11% per 10g ethanol consumed per day was identified.[24] In analyses stratified by hormone receptor status, a meta-analysis of six studies did not find an association with ER-/PR- breast cancer.[24] However, for every 10g in ethanol consumption per day, increased risks of 6% and 12% were seen for ER+/PR+ and ER+/PR-, respectively.[24]

Animal and cell culture models provide evidence that ethanol metabolites enhance mammographic carcinogenesis.[24,135] It has been suggested that derivatives of alcohol act as a carcinogen, increasing DNA damage in breast tissue.[135] Alcohol also may promote the movement of other carcinogens into cells within the breast due to its ability to act as a solvent for other molecules.[24] Characteristics of the diets of high alcohol consumers are likely to contribute to the development of cancer, as they are typically deprived of certain essential nutrients that can subsequently increase the susceptibility of cells to the effects of carcinogens.[24] Based on the CUP's summation of observational studies and the aforementioned biologic plausibility, the WCRF/AICR has concluded that alcohol intake

has a convincing positive association with breast cancer risk, including sufficient evidence of a dose-response relationship, although no threshold in risk has been identified.[24] Furthermore, the previously described epidemiologic evidence of the estrogenic properties of alcohol intake in section 2.2.1 support the hypothesis of estrogen metabolism mediating the association between alcohol intake and breast cancer.

Evidence of an inverse association between dietary fiber and estrogen metabolism, outlined in section 2.2.1, has supported the hypothesis of an association between dietary fiber and breast cancer risk. Data from case-control studies have reported a reduction in risk with increasing fiber intake, but overall evidence is inconclusive according to the WCRF/AICR.[24] One meta-analysis of 16 prospective studies reported a reduction in risk among the highest consumers compared to the lowest for total dietary fiber (RR: 0.93; 95% CI: 0.89-0.98).[150] Similar reductions in risk were observed for fruit fiber, vegetable fiber, and cereal fiber, but the authors reported no association for insoluble fiber.[150] The associations observed between dietary fiber may be a result of facilitated excretion of estrogen, or it could be due to the high correlation between fiber with fruit and vegetable intake. Currently, the WCRF/AICR has concluded there is limited evidence to suggest an association between fruit and vegetable intake with postmenopausal breast cancer risk.[24] In the 2017 CUP report, however, there is suggestive evidence that non-starchy vegetables are associated with decreased risk of ER- subtypes, only.[24] A pooled analysis of over 35,000 cases showed a 18% decrease in risk of developing ER- subtypes when comparing the highest quintile of non-starchy vegetables intake compared to the lowest.[24] Data from the EPIC study reported no association for fruits, but observed a significant inverse association between vegetable

intake and overall breast cancer when comparing the highest and lowest quintiles (HR: 0.87; 95% CI: 0.80, 0.94), with the strongest association observed for ER-/PR- breast cancer cases (HR: 0.74; 95% CI: 0.57, 0.96).[151] An investigation in to the Italian section of EPIC, identified a significant inverse association comparing the highest and lowest quintiles of consumption for total vegetables (HR: 0.65; 95% CI: 0.53-0.81) and for leafy vegetables (HR: 0.70; 95% CI: 0.57-0.86).[152] A meta-analysis of 15 prospective studies identified a significant association for fruits and vegetables combined (HR: 0.92; 95% CI: 0.86-0.98) but not vegetables alone when comparing the highest and lowest quintiles of intake.[20] A subgroup analysis of postmenopausal women identified an inverse association for fruits only (RR: 0.89; 95% CI: 0.83-0.95), but not for vegetables or their combination.[20] A more recent pooled analysis of nearly 1,000,000 women reported no associations between fruits and vegetables, only fruits, or only vegetables with overall breast cancer.[153] However, when only considering ER- breast cancer cases a significant inverse association between the highest and lowest quintiles of vegetable intake was identified (RR: 0.82; 95% CI: 0.74-0.90).[153] Stronger associations were observed in premenopausal women.[153] One explanation for the inconclusive results is the method of consumption. Fruits are almost always consumed raw, but vegetables are cooked in a variety of ways that may alter the availability of constituents that influence breast cancer risk.[154]

The 2017 CUP report has designated foods high in carotenoids with a "limited – suggestive" association with a decrease in breast cancer risk.[24] A meta-analysis of 9 studies showed an 18% decrease in risk of breast cancer per 100 μg/dL of circulating carotenoids, however the report also cited a meta-analysis of 18 studies that found no

association with dietary beta-carotene.[24] In addition to non-starchy vegetables and foods high in carotenoids, diets high in calcium have "limited – suggestive" designation for an association with breast cancer.[24] Six of seven studies cited in the CUP reported an inverse association with postmenopausal breast cancer.[24] In a dose-response meta-analysis, a 300 mg increase of dietary calcium was associated with a 4% reduction in risk.[24] Although mechanisms are unclear, it likely has to do with the prominent role of calcium in cellular signaling that can influence proliferation and apoptosis.[155]

Many other dietary factors have been deemed to have a "limited – no conclusion" designation with respect to the development of postmenopausal breast cancer in the WCRF/AICR's CUP.[24] Dietary fat has been frequently studied with regard to increasing risk of breast cancer, yet the evidence has been inconclusive.[2] A meta-analyses from over 140 mice studies concluded dietary fat promoted mammary carcinogenesis independent of total energy intake.[156,157] However, results from observational evidence have failed to support the animal models.[158] The 2010 CUP report on breast cancer based its "limited" designation for dietary fat on evidence from 10 cohort studies and 16 case-control studies, with no updates in the 2017 report.[24,159] Separate meta-analyses for the cohort and case-control studies included in the report yielded a non-significant and significant positive association for total dietary fat and postmenopausal breast cancer, respectively.[159] Of six cohort studies investigating percentage of total energy intake from fat, the majority reported a decrease in risk, but one study reported a significant positive association with postmenopausal breast cancer risk.[159] In the WHI Dietary Modification Trial, in which the intervention group was meant to reduce fat intake by 20%, no significant difference in risk of postmenopausal breast cancer was seen after 8 years.[110]

However, among women who consumed at least 36.8% of all energy from fat at baseline, a significant decreased in risk was seen in the intervention group compared to controls (HR: 0.78; 95% CI: 0.64-0.95).[110] According to the authors, the HR estimates and upper bound of the CI lowered when accounting for greater adherence to the intervention, suggesting the presence of an association.[110] If there is a true association between dietary fat and breast cancer risk, it has been proposed that dietary fat may work through an influence on estrogen metabolism.[111]

The bulk of evidence from studies of intakes of different types of fatty acids, rather than total fat, has yielded similar inconclusive results. A cohort of nearly 50,000 women identified no association when examining SFA, MUFA, and PUFA in relation to overall breast cancer risk.[160] However, when only considering women over the age of 50, most of whom were presumably postmenopausal, women in the highest MUFA and PUFA quintile intake experienced less incidence of breast cancer compared to the lowest quintile (HR: 0.45; 95% CI: 0.25-0.99 and HR: 0.54; 95% CI: 0.35-0.85, respectively).[160] Inverse associations between PUFA intake and breast cancer risk have been observed, but results are inconsistent.[161,162] Women enrolled in EPIC who were in the highest consumption quintile of SFA had 13% increased risk of breast cancer compared to the lowest quintile (HR: 1.13, 95% CI: 1.00-1.27; p_{trend}=0.038).[163] Meta-analyses of MUFA and breast cancer have reported both positive[164] and inverse associations with breast cancer risk.[165,166] The sources of the MUFA may be one reason for the inconsistencies.[162] Studies of fat from animal sources in association with breast cancer have also been inconclusive.[23]

Consumption of soy foods has been associated with reduced risk of breast cancer, however results across different study populations have been inconsistent.[167] Two different recent meta-analyses identified 35 and 14 studies investigating an association between soy and breast cancer.[168,169] The former identified a significant inverse association comparing the highest consumption groups to the lowest groups (RR: 0.89; 95% CI: 0.79–0.99).[169] When stratified by the origin of the study population, the association remained significant for Asian countries (RR: 0.76; 95% CI: 0.65–0.86) but not in Western study populations (RR: 0.97; 95% CI: 0.87–1.06).[169] The other analysis based on a smaller number of studies, stratified by menopausal status and reached the same conclusion for both pre- (OR: 0.59, 95%CI: 0.48-0.69) and post-menopausal Asian women (OR: 0.59, 95% CI: 0.44-0.74).[168] In Western populations, results from premenopausal women were not significant and postmenopausal women exhibited an inverse association nearing significance (OR: 0.92; 95% CI: 0.83-1.00).[168] Another literature review concluded there was no association between breast cancer and soy consumption in Japanese women.[170] The differences in the associations observed between Asian and Western study populations is most likely driven by the relative intakes of soy foods, which is much more common among Asian countries.[167] It also has been hypothesized that early life exposure to soy may be more important than intake in adulthood.[132,167] As mentioned in section 2.2.2, it is possible genetic polymorphisms affect the relationship between soy intake and breast cancer risk, through modulation of soy's effect on estrogen metabolism.[124]

2.3.2 Dietary patterns and breast cancer

As shown in section 2.3.1, for many nutrients and dietary components there is inconclusive evidence of an association with breast cancer risk. It is possible the uncertainty in the hypothesized relationship between breast cancer and diet is due to the complex interactions that occur in reality when combinations of foods and nutrients are consumed. The USDA's DGA called for a focus on dietary patterns because "the totality of diet […] may have synergistic and cumulative effects on health and disease."[171] Dietary pattern analyses incorporate the potential for this web of influence by assessing diet in its entirety, accounting for multiple foods consumed, rather than singular specific components. Therefore, dietary pattern analyses may detect a dietary effect on breast cancer due to the combinations of foods, that is not seen when studying isolated components. However, similar to single nutritional factors, the evidence of an association between dietary patterns and breast cancer has been inconclusive.[17,18,21,39,40,172]

There are two prevailing methods used in dietary pattern analyses.[173] Data-driven patterns, or *a posteriori*, are empirically determined from each study population in which the analysis occurs.[26,173,174] Within data-driven patterns, methods can be further delineated by the outcome-dependent or -independent properties of the approach.[26,173,174] Contrastingly, investigator-defined patterns, or *a priori*, are based on hypotheses of diet-disease relationships or on certain guidelines that constitute a healthy diet, before any analysis occurs.[26,173,174] It is possible that the inconsistency of associations between dietary patterns and breast cancer is a result of the high heterogeneity in applied methodologies to derive and study dietary patterns.

In studies of breast cancer and data-driven dietary patterns, the "Western" or "unhealthy" pattern is often hypothesized to increase the risk of breast cancer because it is typically characterized by high intakes of animal products, refined grains, and sugars. One meta-analysis did not identify an association,[18] however multiple reviews and original research articles have supported evidence for a positive association with breast cancer.[17,18,36,175–178] Among studies reporting no association between the "Western" pattern and overall breast cancer, multiple studies identified a significant positive relationship when limiting to postmenopausal,[36,177–179] ER+,[177,180] or normal weight women.[180] The "prudent" or "healthy" dietary pattern with high intakes of fruit and vegetables, whole-grains, legumes, olive oil and fish, has shown a more consistent association with evidence of a reduction in risk.[17,18,21,28,177,181] Similar to the "Western" pattern, some studies suggest the magnitude of the association is strongest in postmenopausal[21,182] or normal weight women,[38,183] or with ER+ subtypes.[37] When looking at only vegetarians and non-vegetarians, no significant difference in risk has been suggested.[123,184] Using data from the EPIC-Potsdam study, a dietary pattern was derived using RRR to explain variation in fatty acid intake (SFA, MUFA, ω-3 PUFA, ω-6 PUFA).[185] Women in the highest tertile of the pattern had twice (HR: 2.00; 95% CI: 1.30 – 3.09) the risk of developing breast cancer, with no effect modification by menopausal status.[185]

The heterogeneity of food groups identified in the "Western" or "prudent" patterns is shared in development of *a priori* patterns. Although they differ in what constitutes a healthy diet or are aimed at prevention of different diseases, some evidence of associations with breast cancer have been shown. Typically, higher scores on these *a*

priori patterns correlate with high intakes of fruits, vegetables, legumes, whole grains, and seafood. Contrastingly, low scores correlate with high intakes of red meat, highly processed foods, including refined grains, and other animal products. In addition to the AHEI, common *a priori* dietary patterns are the Recommended Food Score (RFS) which is based on current intake guidelines in the U.S., and the Diet Quality Index Revised (DQI-R) from the National Research Council. AHEI, which was inversely associated with estrogen, also was inversely associated with postmenopausal breast cancer (HR: 0.78;95% CI: 0.59-1.04; p_{trend}=0.01) when comparing across quintiles, but not for overall breast cancer.[186] Another study also reported no association between the AHEI and overall breast cancer risk, however they did not stratify by menopausal status.[187] Similarly, the association for the RFS (HR: 0.69; 95% CI: 0.51-0.94) with postmenopausal breast cancer was only present among ER- cases.[188] Although they were not statistically significant, associations with the AHEI and RFS with overall postmenopausal breast cancer showed a consistent inverse association.[186–188] The DQI-R was not associated with breast cancer except women with genetic predispositions to breast cancer.[187,188] Although the majority of studies using dietary patterns based on guidelines failed to find statistically significant associations, all have shown inverse associations, suggesting overall diet quality may reduce the risk of breast cancer. Selection of foods that are hypothesized to have an effect on breast cancer in an *a priori* pattern, and not necessarily foods that constitute an overall healthy diet, may result in stronger associations.

Like the AHEI, the previously described MeD has exhibited anti-estrogenic effects,[147] and is inversely associated with many chronic diseases.[189] One recent review

reported weak evidence of an association between MeD and breast cancer from observational studies.[40] However, a meta-analysis of 23 observational studies reported an inverse association (RR: 0.93; 95%: CI 0.87-0.99).[190] In the NHS, the aMeD was only statistically significant for ER- breast cancer among postmenopausal women comparing the highest quintile to the lowest (HR: 0.79; 95% CI: 0.60-1.03; p_{trend}=0.03).[188] In a randomized controlled trial of over 4,000 women aged 60 to 80, women allocated to the MeD supplemented with extra virgin olive oil intervention group experienced nearly 70% less risk of breast cancer than the control group (HR: 0.32; 95% CI: 0.13-0.79).[48]

Another commonly used category of *a priori* dietary patterns is those developed based on hypothesized disease pathways. The DASH diet was developed as a potential tool for intervening on hypertension.[191] Surprisingly, even though it was developed based on a mechanistic pathway for a different disease, the DASH diet has shown an association with breast cancer, although only for HER-2 positive cases (HR: 0.44; 95% CI: 0.25-0.77).[186] A dietary pattern developed on the basis of foods associated with inflammatory markers, the Dietary Inflammatory Index (DII™), has shown mixed results for breast cancer.[49–51,192,193] Although one study reported no association with postmenopausal breast cancer,[192] others have reported significant associations with breast cancer, with larger estimates observed in postmenopausal[49,193] or obese women,[49] and with breast cancer mortality.[51] Together, the evidence suggests a pro-inflammatory diet is associated with greater incidence and mortality from breast cancer. Specific to breast cancer, a previously described pattern was based on food groups correlated with circulating estrogen levels, which was subsequently not associated with postmenopausal breast cancer in NHS.[32] However, when the same pattern was applied in a Swedish

cohort, a 29% increase in breast cancer risk (HR: 1.29; 95% CI: 1.08-1.55) was reported when comparing women in the highest quartile with the lowest.[27]

2.4 Risk factors for breast cancer

Breast cancer is the most diagnosed cancer among women after non-melanoma skin cancer, with over two-thirds of cases occurring in women over the age of 55, and results in the second most cancer-fatalities after lung cancer.[2] Established risk factors for breast cancer include age, obesity, physical inactivity, alcohol intake, and reproductive factors, most of which affect the development of mammographic tumors through hormonal influences. Due to the high incidence of breast cancer, modifiable primary prevention methods, such as dietary intervention, are of great interest.

2.4.1 Menopausal status

There is evidence that risk factors, incidence, and prognosis of breast cancer vary between pre- and postmenopausal women, highlighting the significant differences between the two disease strata.[2,194,195] The heterogeneity in the two diseases may be crucial to explaining some of the inconclusive findings in the relationship between diet and breast cancer, as the grouping of both menopause statuses as one occurs frequently in the literature.[21] The onset of menopause is a marker for a reduction in ovarian endocrine activity. Subsequently, levels of sex hormones, including estrogen, are significantly attenuated in postmenopausal women.[56,58] Sex hormones in premenopausal women have high within-person variability corresponding to their menstrual cycle.[196] Contrastingly, due the termination of menstruation after menopause, postmenopausal women have lower, less variable levels of circulation estrogen.[56,197] The reduced variability and

magnitude of the hormone levels is hypothesized to make postmenopausal women more sensitive to estrogenic effects in relation to breast cancer risk.[4,9,195,198,199] This hypothesis is supported by many of the estrogen-related risk factors for breast cancer, as described below, which appear to have a greater effect in postmenopausal women.[2,13,195] In addition to, and partially as a result of the hormonal changes after menopause, there are paralleled atrophic changes to mammary tissue, with increasing amounts of adipose in the breast.[198] Increased amounts of adiposity in the breast results in higher localized levels of estrogen as a result of the estrogenic properties of adipose tissue.[198]

2.4.2 Weight status and physical activity

Multiple factors related to increased adiposity and PA are associated with the development of breast cancer in postmenopausal women.[13,200,201] Using the Centers for Disease Control and Prevention's (CDC) BMI cutoffs for overweight (25.0-29.9 kg/m^2) and obesity (≥30 kg/m^2), risk of postmenopausal breast cancer is 1.5 and 2 times that of normal weight women (18.5-24.9 kg/m^2), respectively.[2] In the WCRF/AICR's Second Expert Report CUP, total body fatness has "convincing" evidence and biological plausibility to increase risk of postmenopausal breast cancer.[24] The designation is based on an updated meta-analysis of more than 56 studies showing a 12% significant increase in risk per 5 kg/m^2 increase in BMI, with stronger evidence among ER+ subtypes.[24] In addition to total body fatness, measures of abdominal fatness, such as waist circumference (WC) and waist-to-hip ratio (WHR) have a "probable" association with increased risk of postmenopausal breast cancer.[13,24] Pooled evidence from 11 cohort studies showed an 11% increase in risk for an 10 cm increase in WC.[24] Similarly, the report cited a 10% increase for a 0.1 increase in WHR.[24] Estimates were slightly

attenuated, but still significant, when only considering studies that adjusted for BMI.[24] Women who gain weight as adults are even more susceptible to breast cancer.[2,13] For a 5 kg gain in weight during adulthood, a meta-analyses of 15 studies reported a 6% significant increase in risk of postmenopausal breast cancer.[24]

The biologic mechanisms of increased weight status and postmenopausal breast cancer risk are due to the hormonal properties of adipose tissue.[2,24] The chronic state of inflammation that is present in obese individuals is mediated by adipokines, such as tumor necrosis factor alpha (TNF-α).[53,202] The downstream effects of adipokine secretion lead to an altered immune response that can facilitate cell proliferation and tumor growth.[203] Furthermore, TNF-α in adipocytes inhibits glucose uptake resulting in sustained levels of increased insulin.[204,205] There is some evidence hyperinsulinemia is associated with increased breast cancer risk, likely due to its ability to promote DNA synthesis and the activity of insulin-like growth factor (IGF).[63] The influence of IGF on breast cancer risk has become increasingly apparent, primarily due to its mitogenic properties affecting cellular growth and differentiation.[63]

The predominant hypothesis by which increased weight status, specifically increased accumulation of adipose tissue, affects postmenopausal breast cancer risk is the ability of adipose tissue to synthesize estrogen.[62,63,206] Adipose tissue is the largest source of endogenous estrogen in postmenopausal women, and there is strong evidence for a positive linear association between adipose tissue and estrogen levels in postmenopausal women.[2,63,207] Adipose tissue contains high levels of the enzyme aromatase, which plays a significant catalytic role in estrogen synthesis.[63] Aromatization is the last step in the conversion of cholesterol to estrogen for both estradiol and estrone.[53] The influence of

adipose tissue on breast cancer risk through estrogen metabolism is evident when looking at various strata of estrogen-related breast cancer risk factors. For example, the influence of HRT on breast cancer risk is strongest among lean women, likely because the exogenous estrogen from the therapy has a relatively greater effect in the absence of (or reduced amount of) adipose-derived estrogen.[208]

Contrary to postmenopausal breast cancer, increased adiposity is associated with a decrease in the risk of breast cancer in premenopausal women.[24] Results from the WCRF/AICR's CUP meta-analysis indicated an 7% decrease in risk of premenopausal breast cancer for every 5 kg/m^2 increase in BMI.[24] The mechanisms behind the inverse association between adiposity and premenopausal breast cancer are unclear. It has been hypothesized that the increased levels of adipose-derived hormones, such as IGF, may promote anovulation which reduces a woman's lifetime exposure to estrogen.[209] It is also possible that increases early life exposure to adipose-derived estrogen may alter breast differentiation in a way that is beneficial to prevent malignancies.[210]

Potentially through its effects on adiposity and other mechanisms, energy expenditure through PA has an inverse association with postmenopausal breast cancer risk.[2] The majority of cohort studies in the WCRF/AICR's report showed a significant inverse association between recreational PA and postmenopausal breast cancer, resulting in a "probable" designation for decreasing risk.[24] A meta-analysis yielded a 13% reduction in risk when comparing the highest level of PA with the lowest, with a similar 10% reduction in risk when only looking at vigorous PA.[24] The hypothesized beneficial effect of increasing PA is related to the promotion of metabolic efficiency, translating to a reduction of adipose tissue and increase in lean mass.[62] Subsequently, PA improves

insulin response and protects against chronic inflammation.[62] Increased PA has also been shown to have an inverse association with circulating estrogen, possibly through increased levels of SHBG.[24,62] Intervention studies have shown a reduction of circulating estrogen after participating in PA, suggesting PA may reduce breast cancer risk through attenuation of exposure to estrogen.[211,212]

2.4.3 Hormone replacement therapy and contraceptives

Exogenous hormones use, such as in contraception or postmenopausal HRT, has a positive association with breast cancer incidence.[2] There is evidence that use of OCs that contain estrogen and/or progesterone, has a minor effect on risk.[2] Women who use OCs, specifically those manufactured with high hormonal dosage, show the greatest increase in risk when use starts before the age of 20.[2] The increase in risk attenuates when use of OCs is terminated.[2] Evidence suggests that a previous user of OCs has the same risk profile of someone who never used if it has been at least 10 years since their last use.[2]

The other main source of exogenous estrogen, HRT, is used among women who underwent hysterectomy and cannot produce their own estrogen, or among women who are trying to mitigate the effects of menopause due to low levels of estrogen. There has been a drastic reduction in the latter HRT use after initial results from the landmark randomized controlled trial in the WHI.[213,214] Originally designed to investigate a hypothesized protective effect of estrogen plus progestin in relation to coronary heart disease and all-cause mortality among women, the trial was prematurely terminated as intermediate results identified an increase in risk of many conditions in the intervention group.[214] Compared to controls, women who underwent estrogen plus progestin therapy had a significant 26% increase in risk of breast cancer.[214] Increasing duration of use

showed stronger associations, however, termination of use causes a women's risk to revert to what it would be if she never used, similar to what has been observed among OC users.[214,215] Results from the WHI were corroborated in multiple other studies with regards to the effect of estrogen plus progestin.[208,215,216] Interestingly, there is evidence of an effect modification by BMI in the association between HRT and breast cancer risk.[208] Although adipose tissue promotes estrogen production, the risk estimates between HRT and breast cancer were higher in lean women, compared to obese women, in a reanalysis of 51 observational studies.[208] It is possible that the amount of estrogen produced by adipose tissue is enough to cause a sufficient increase in risk, thereby masking any additional effect of HRT use on breast cancer risk. This would explain why HRT has a greater effect on risk among lean women, because these women do not have as much adipose-derived estrogen.

There are forms of HRT that do not use the combination of estrogen plus progestin, which have shown inconclusive results regarding breast cancer risk.[208] The Million Women Study in the United Kingdom showed a significant 30% increased risk of breast cancer among women who used an estrogen-only replacement therapy compared to women with no HRT.[215] Contrastingly, women in the WHI's estrogen-only trial showed evidence of a significant decrease in risk after adjustment for adherence (HR: 0.67; 95%CI: 0.47-0.97).[217] Tibolone, a synthetic hormone with androgenic properties, has also been shown to increase risk.[215]

2.4.4 Reproductive factors

Numerous factors regarding a women's reproductive history can influence their risk of postmenopausal breast cancer, often also effecting their exposure to endogenous

estrogens, as evident by a stronger association with ER+ cases.[2] The earliest reproductive factor affecting breast cancer is age at menarche. There is an inverse association between age at menarche and breast cancer, with women who experience menarche at the age of 12 and younger with the greatest risk.[2,218] After the age of 12, a 10-20% reduction in risk has been estimated for each 1-year increase in age that menarche occurs.[218] Similar to the relationship with late onset of menopause and breast cancer, women who experience early menarche typically have a greater lifetime exposure to ovarian hormones.[2,218]

Once of child-bearing age, those women who never have children, or do so at an older age, are at an increased risk of breast cancer compared to women who have an earlier age at first birth.[2,219,220] Increasing parity and age at first birth are both inversely associated with breast cancer.[2,219,220] Compared to nulliparous women, those who were parous have significantly lower levels of serum estrogen and greater concentrations of SHBG. Therefore, it is plausible that multiparous women who gave birth at a young age have a lower lifetime exposure to estrogen.[221] However, it also is possible that a woman's nulliparity status results from infertility due to low levels of steroid hormones, which would indicate a lower exposure to estrogen.[222]

Among women who have children, there is evidence that those who breastfeed are at a lower risk.[2] Furthermore, the longer a women breastfeeds has shown greater reduction in risk.[2] The mechanism behind this decrease in risk most likely has to do with increased differentiation of breast tissue, however it is also possible the paralleled inhibition of menstruation that occurs during lactation plays a role.[2,13] By inhibiting the number of menstrual cycles, lactation reduces a woman's lifetime exposure to endogenous estrogen.[2,13]

2.4.5 Inherited risk

Although most incident breast cancer cases occur in women without a history of the disease, there is a strong link between risk and an individual's personal and family history of breast cancer. Early onset breast cancer is often a result of inherited risk, as genetic factors are likely to have a stronger influence, whereas accumulation of environmental and lifestyle factors take effect in cases of older, postmenopausal women.[2,223] Women who have one, two or at least three first-degree relatives experience 2, 3, and 4 times the risk of developing breast cancer, respectively.[2] The younger the relatives were diagnosed the stronger the association with risk in family members.[2] In addition to family history of breast cancer, women with relatives who have been diagnosed with ovarian, prostate and endometrial cancers, all of which are cancers with strong hormonal properties, are also at increased risk.[2,224] Women who have previously been diagnosed with cancer are approximately 1.5 times as likely to develop a secondary breast cancer compared to women with no personal history.[2]

For some previously diagnosed women, the increased risk of secondary breast cancer is due to their genetic predisposition. The strongest associated and most frequently analyzed genetic mutations for increased breast cancer risk are in the *BRCA1* and *BRCA2* genes.[2] Although the mutations occur in less than 1% of the female population, there are estimates that they account for as much as 10% of all breast cancer cases.[2] Women who carry the a mutation in *BRCA1* and *BRCA2* have between a 50-80% lifetime risk of breast cancer, compared to a 12% lifetime risk in the general population.[225] There is evidence that other genetic variations present low increases in risk, in addition to a strong belief

that these variations interact with lifestyle factors, such as dietary habits, to affect breast cancer risk.[2,226]

2.4.6 Demographics

As with most major chronic diseases, there are multiple demographic risk factors strongly associated with incident breast cancer. Incidence rates differ among many strata of age, social class, ethnicities, and races. Many, but not all, demographic and socioeconomic risk factors for breast cancer are related to screening behaviors.[2] For example, the lifetime risk of developing breast cancer has increased in the past 40 years, partly as a result of increased life expectancy, but also due to better detection and increased participation in screening.[2]

The strongest risk factor for breast cancer is age, due to the prominent role of cellular damage, or mutations, in the development of proliferation of cancer cells.[6] As women age and the number of cellular divisions take place over time, there is a greater chance of improper division and damage to the DNA. The subsequent effect of the damage in the mutated DNA is exacerbated in the diminished capability of cellular repair mechanisms of older individuals.[6] Furthermore, environmental exposures that accumulate over time can result in DNA damage and alter DNA expression.[227] According to the American Cancer Society (ACS), the median age at breast cancer diagnosis in the U.S. was 61 years old between 2008-2012.[2] The Surveillance, Epidemiology, and End Results (SEER) program estimated the age group with the highest percentage of incident cases is between 55-64 years old using data from 2009-2013.[228] Only 10.7% of all new cases occur in women under the age of 45, whereas 68% of all new cases occur in women 55 years and older.[228]

Annual age-adjusted incidence rates reported in the 2009-2013 SEER database were highest among White women (128.0 per 100,000), with Black women experiencing similar rates (125.2) during this time period.[228] However, comparing White and Black women, the ACS reported significantly higher rates of breast cancer among White women between the ages of 60 to 84, and higher rates in Black women younger than 45.[2] In addition to being diagnosed at younger ages, Black women are more likely to have aggressive cases, such as triple negative, or advanced stage cancer and subsequently higher breast cancer mortality over their lifetime compared to White women.[228,229] Risk of developing breast cancer is lower among Hispanics and Asian/Pacific Islanders compared to Black and White women.[2,228]

Regardless of ethnicity or race, socioeconomic status has repeatedly shown a positive association with breast cancer incidence, using education, income, or their aggregate measure to define social class.[229–231] More years of education and highest degree obtained have both shown positive associations with breast cancer incidence,[232–234] as well as annual income[234,235] and occupational supervisory rank.[236] This association is strongly influenced by screening behaviors, as shown in ACS data from 2010 where the prevalence of a mammography within the past two years ranged from 24-28% less in poor women (defined as 100-199% of poverty) compared to non-poor.[206] Incidence is lower and mortality is higher among women who reside in rural areas compared to urban dwellers, due to the aforementioned reduced access to screening and detection at a more advanced stage.[237–239]

2.4.7 Tobacco use

Evidence of an association between smoking tobacco and breast cancer has been suggestive but inconclusive.[2] Some have hypothesized an association among those who are heavy smokers, or who have been smoking for a long duration.[2] A recent meta-analysis reported an 8% increase in risk of breast cancer when comparing current smokers and never smokers using data from 27 prospective studies (RR: 1.08; 95 % CI: 1.02-1.14).[240] When looking at passive smoking, a meta-analysis reported an increase in risk (OR: 1.62; 95% CI: 1.39-1.85) but no association with active smoking.[241] Together, the results suggest tobacco smoke may play a role in developing breast cancer. The predominant pathway by which tobacco smoking affects breast cancer is through increased inflammation, along with the carcinogenic effects of tobacco smoke.[242] Some have identified associations between high levels of estrogen and smoking,[12] while others have reported an anti-estrogenic effect,[243] and even associations with the 2-hydroxylation pathway suggesting a beneficial alteration of estrogen metabolism with smoking.[244]

2.4.7 Lifestyle indices

There is evidence of an association between individual modifiable lifestyle characteristics, such as PA and alcohol use, with development of postmenopausal breast cancer. Lifestyle factors often cluster together in individuals who adopt healthy or unhealthy lifestyle, so it may be beneficial to study lifestyle factors using a combined lifestyle score.[245] An *a priori* healthy lifestyle index score (HLIS) based on diet, tobacco use, alcohol, PA and BMI reported 21% lower risk of breast cancer (HR: 0.79, 95% CI: 0.73-0.85) among the fourth, or most healthy group, compared to the second group in the EPIC cohort.[246] Application of the HLIS, with a slight modification of the diet to include

fish, folate, glycemic index, and other breast cancer risk-specific dietary components also showed an inverse association with postmenopausal breast cancer risk comparing the highest category to the second (HR: 0.74; 95% CI: 0.66–0.83).[16] The association was strongest for ER-/PR- (HR: 0.60; 95% CI: 0.40-0.90) but also significant for ER+/+ breast cancer (HR: 0.81; 95% CI: 0.67-0.98).[16] In both of the previously mentioned HLIS's, the second group served as the referent due to low numbers of individuals adopting the healthiest behaviors for some of the scoring components in the first group. Also using data from EPIC, a lifestyle score was developed to evaluate adherence to the WCRF/AICR recommendations on body fatness, PA, energy dense foods and drinks, plant foods, animal foods, alcohol use, and breastfeeding in women. Compared to the lowest scores, all categories showed a significant inverse association with breast cancer, with the strongest association in the highest scoring groups (HR: 0.84; 95% CI: 0.78-0.90).[14] Adherence to WCRF/AICR recommendations and their association with breast cancer risk has been studied in other populations, as well.[247–249] In the Swedish Mammography cohort, women who met at least six of the seven recommendations had nearly half the risk (HR: 0.49; 95% CI: 0.35-0.70), with a greater reduction in ER+/PR+ subtypes compared to ER-/PR-.[247] In the Iowa Women's Health Study, an inverse association was observed with postmenopausal breast cancer, and that association did not differ in the presence of non-modifiable risk factors, such as taller height, family history of breast cancer, or greater number of potentially fertile years.[249]

Some lifestyle scores have been developed for a specific study population. One score was developed to assess increasing incidence of breast cancer among indigenous women in New Zealand using 11 scoring criteria (red meat, protein, seafood, energy

dense foods, solid fats, plant foods, smoking, exercise, BMI, and breastfeeding).[250] No association was observed among non-indigenous women but the highest lifestyle score tertile had a significantly lower odds of breast cancer (OR: 0.47; 95% CI: 0.23-0.94) compared to the lowest tertile among indigenous women.[250] Investigators of a large case-control study of Mexican women developed a similar lifestyle score using the same five components, except adherence to the "Western" diet was used to inversely derive the dietary component.[15] The authors reported an inverse association with breast cancer in postmenopausal women (OR: 0.20, 95% CI: 0.11–0.37) when comparing the highest versus lowest quintiles, with PA and alcohol use as the main contributors to the association.

CHAPTER 3

METHODS

3.1 Statement of aims and hypotheses

The overarching goal the dissertation work was to derive and evaluate a dietary pattern based on estrogen metabolism in relation to postmenopausal breast cancer risk solely as a dietary exposure and as part of an aggregate score for estrogen-related lifestyle factors. We hypothesized that a dietary pattern that is characteristic of increased estrogen exposure would be positively associated with postmenopausal breast cancer. An aggregate lifestyle score representative of habits that are beneficial to estrogen metabolism was hypothesized to be inversely associated with postmenopausal breast cancer. We expected to see the strongest associations for ER+ cancer subtypes, with effect modification by other estrogen-related risk factors for breast cancer.

In Aim #1, a dietary pattern was developed based on food groups associated with various measures of estrogen metabolism, and was subsequently applied in a prospective investigation into postmenopausal breast cancer risk. We hypothesized that diets high in animal products, and low in vegetables and fiber would be associated with high estrogenic potential, measured as a high ERDP score. Similar to the first aim, we investigated an association between the ERDP with postmenopausal breast cancer risk in Aim #2, but used a study population different from the one in which it was derived. In both prospective investigations, using the PLCO and SS, we hypothesized a positive association between the ERDP and incident breast cancer. We expected to see a stronger

association among breast cancer cases that are ER+ compared to ER-. We also hypothesized that the strongest association would be observed in strata of effect modifiers assumed to lower estrogen exposure, such as leaner women compared to overweight women, where the estrogenic effect of diet will have a larger relative influence. In Aim #3, the ERDP was incorporated into the ERLS with alcohol intake, BMI, and PA, all of which are hypothesized to influence estrogen metabolism. We hypothesized that higher ERLS scores, representative of a lower collective estrogenic effect of lifestyle factors, would be inversely associated with postmenopausal breast cancer. Like the ERDP, we expected to see the largest magnitude of associations for ER+ cases, and among strata of effect modifiers that have a smaller estrogenic effect.

3.2 Descriptions of the study populations

Multiple study populations, including a subset of one of the larger studies, were used to complete the dissertation aims. Participants of PLCO were utilized in Aim #1 and Aim #3. Derivation of the ERDP within Aim #1 took place in a subset of PLCO participants with information on baseline serum EM concentrations, which is described in detail after an overview of PLCO below. To examine the ERDP in a study population external to the one in which it was developed, the SS was used in Aim #2. Detailed descriptions can be found in the next three sections.

3.2.1 Prostate, Lung, Colorectal and Ovarian Screening Trial

An initiative of the NCI, the PLCO is a large population-based screening trial designed to determine the effects of screening on cancer prognosis and mortality. Design and implementation has been described in detail elsewhere.[33] Briefly, participants were

recruited between 1993 and 2001, with the intervention trial completing in 2006 and follow up continuing through 2015. Recruitment of 76,685 men and 78,216 women aged 55 to 74 at enrollment took place at 10 different screening centers across the nation. After randomization to the intervention arm, women participated in regular chest x-rays, flexible sigmoidoscopy, CA-125 blood tests, and transvaginal ultrasound during the first six years and were followed up for an additional seven years. Women were excluded at recruitment if they had a history of lung, colorectal, or ovarian cancer. If women were currently undergoing treatment for any previously diagnosed cancers, or if they were participating in another screening or primary prevention trial, they were also excluded. Prior to October 1996, women who previously had both ovaries surgically removed were excluded from enrollment. Eligible participants underwent a physical examination and filled out a questionnaire with information on demographics, medical history, family history, lifestyle factors, and recent history of participation in screening examinations at baseline.

For the dissertation work, only data from the 39,104 women randomized to the intervention arm of the study, who participated in standard of care screening practices, were used. Use of only women in the intervention arm is required for a couple of reasons. First, the sample of women included in the nested case-control study with baseline serum estrogen data were selected as a subset of the intervention arm. Additionally, only women in the intervention arm were asked to complete the dietary questionnaire (DQX) at baseline. A different dietary instrument, the diet history questionnaire (DHQ), began to be administered to both arms of the study 3 years after baseline. Therefore, it would be inappropriate to use the DHQ in an investigation of baseline serum estrogen levels due to

the issue of temporality. Before any analytic exclusions were made, less than 15% of women in the intervention arm self-identified as a racial/ethnic minority: 5.5% non-Hispanic Black (n=2,170), 1.5% Hispanic (n=605), and 3.2% Asian (n=1,259).

3.2.1 PLCO nested case-control

A subset of postmenopausal women randomized to the intervention arm of PLCO with information on serum EMs were included in the analyses to derive the ERDP. Complete information on the nested study has been published elsewhere.[35] Briefly, the nested study population was drawn from all 1,141 incident breast cancer cases diagnosed from the start of recruitment in 1993 through June 30, 2005, and a random sample of 1,141 control subjects. After excluding women who were not postmenopausal, were using HRT at baseline, or had prior diagnoses of cancer, the sample was reduced to 390 cases and 453 controls. For the purposes of the present analysis, cases who were diagnosed <2 years after serum sample donation (n=98) were excluded to avoid the possibility of disease processes affecting estrogen levels. Women without a valid DQX (n=77) or with implausible EM levels (i.e., if they were outside of 25th and 75th percentile, plus/minus three times interquartile range; n=15) were further excluded. The final analytic sample included 393 controls and 260 confirmed cases, with a mean of 5.25 years from sample donation to breast cancer diagnosis among cases. Use of some cases served to increase the sample size for this aim and is justifiable because most breast cancer cases are diagnosed without symptoms. Therefore, we believe their diets likely did not change dramatically leading up to their diagnosis. Details of the laboratory methods used are explained in section 3.3.

3.2.3 Sister Study

The SS is a large prospective cohort study designed by the NIEHS to investigate environmental and genetic determinants of breast cancer.[34] A total of 50,884 women aged 35 to 74 who had a sister that was diagnosed with breast cancer were recruited between 2003 and 2009 from all 50 U.S. states and Puerto Rico. Community based recruitment efforts were used through local volunteers, study participants, local and national events, and extensive media campaigns. Recruitment strategies included attempts to enroll women who were of racial/ethnic minorities, older age, and lower income. After enrollment, baseline information was collected via Computer Aided Telephone Interview (CATI) and self-completed risk factor questionnaires on demographics, dietary information, lifestyle and medical history, and exposures from the prior 24 hours. Study staff conducted a home visit to collect blood and urine samples, toenail clippings, and dust collection from the home for environmental exposures. Anthropometric and blood pressure measurements also took place at the home visit. All women are being followed up for at least 10 years. Participants were contacted annually for brief health updates, with a comprehensive follow-up questionnaire administered every two to three years.

A similar proportion of participants enrolled in SS identified as non-Hispanic White (81.0%; n=42,558) as in the PLCO study population. The distribution of racial/ethnic minorities was slightly different, however, with 8.5% identifying as non-Hispanic Blacks (n=4,462) and 4.8% identifying as Hispanic (n=2,515). Over half of the participants were aged 55 and older at baseline (50.97%). In the present work, all women who contributed person-time after the onset of menopause with complete information were used in the investigation of the ERDP and breast cancer.

3.3 Dietary assessment

In PLCO, usual dietary intakes over the prior 12 months were collected via the DQX, which is a 137-item food frequency questionnaire (FFQ) administered at baseline. Over 82% of participants in the intervention arm completed the DQX. Dietary data in SS was assessed using two different versions of the 110-item 1998 Block full diet FFQ. Version 2 of the FFQ contains the same information as the first version, with additional questions on organic foods, microwave use, restaurant and frozen foods – none of which were used in the current dissertation work. Over 16% of SS participants completed version 1, with another 81% completing version 2, totaling over 97% of participants with dietary information.

A qualitative side-by-side comparison of dietary assessment tools used in PLCO and SS showed strong agreement in the foods measured. Both study populations filled out dietary information on usual food consumption, preparation methods, and supplement use. Overall, both dietary assessment tools include the same foods with some minor differences. The DQX had a larger number of line items dedicated to fruits and vegetables than the Block FFQ. However, most of the fruits and vegetables in the DQX are on the 1998 Block FFQ, just combined into fewer lines. Other minor differences included a greater number of lines designed to assess grain intakes and nuts/seed consumption on the 1998 Block FFQ. Both the DQX and 1998 Block FFQ used the USDA's National Nutrient Database for Standard Reference for nutrient analysis.

Usual intakes from the dietary assessment tools were categorized into 1 of 29 food groups based on the USDA's My Pyramid Equivalents Database (MPED).[251] Additional groups were added for cruciferous vegetables, coffee and tea because of their

suggested influence on estrogen metabolism or breast cancer. A number of other groups were omitted to reduce the redundancy of some commonly eaten foods. Namely, the "total" groups for each section was excluded. For example, the "total fruits" group was removed because of the foods that would be contained in the "total fruits" group are accounted for in the "citrus fruits, melons, and berries" or the "other fruits" groups. In total, 32 food groups were used. The food groups, presented in **Table 3.1**, were used as the predictor variables in the RRR analysis, which is explained in greater detail in section 3.6.1.

3.4 Estrogen metabolite measurement

Serum samples collected at baseline and stored at −80°C from women in the PLCO nested study were thawed at 4°C. The previously described LC/MS-MS assay was used to concurrently quantify levels of 15 EMs. The parent estrogens were measured along with their metabolites in the 2-, 4-, and 6-hydroxylation pathways make up the 15 EM (estrone, estradiol, 2OHE-1, 2-methoxyestrone, 2-hydroxyestradiol, 2-methoxyestradiol, 2-hydroxyestrone-3-methyl ether, 4OE-1, 4-methoxyestrone, 4-methoxyestradiol, 16OHE-1, estriol, 17-epiestriol, 16-ketoestradiol, and 16-epiestriol). Quantification of the individual metabolic pathways allows for ratios of those pathways to be used, which is potentially influential in the development of postmenopausal breast cancer.[11] The specifics of sample preparation and LC/MS-MS methods have been described in greater detail elsewhere.[41] An enzyme hydrolytic step is used to elucidate the unconjugated and conjugated forms of parent estrogens. Quality control was assessed using four samples which were inserted into each batch by blinded laboratory staff. The

CV for all EMs was <5%, with even lower CV evident for the parent estrogens (<3%) and unconjugated estradiol (<2%).[35] Levels of EMs between 1–2 pmol/L were able to be quantified in this population of postmenopausal women, with no EMs in the study at undetectable readings.[35]

3.5 Breast cancer ascertainment

Incident breast cancer cases among postmenopausal women in PLCO were identified primarily through self-report via annually mailed follow-up questionnaires, or through the National Death Index, physician reports, state cancer registries, and next of kin reports. Over 96% of the cases were confirmed through hospital records.[252] Using most recent follow-up data from PLCO, a total of 1,652 cases of breast cancer have been ascertained over an average follow-up of about 11.5 years, with 1,316 of the cases diagnosed as invasive (before analytic exclusions). Prior to 2007, breast cancer cases in PLCO were confirmed from medical records with only information on diagnosis date and codes from the second edition of the International Classification of Diseases for Oncology. After 2007, a Breast Cancer Supplemental form was used to capture more information, include ER status of the tumor. There was limited data on ER status of *in situ* cases before the implementation of the supplemental form, as ER status was not routinely assessed among *in situ* cases in the past. The supplemental form was available for 98% of the cases. ER status was available for 70% of total cases (75% of invasive and 35% of *in situ* cases).

In the SS, incident breast cancer cases were ascertained via completion of annual health updates, biennial surveys, and the National Death Index. Response rates for the

surveys were over 94%.[253] Medical record abstraction was used to confirm over 80% of cases and to identify information on treatment and diagnosis, such as ER subtype.[254] Agreement between self-reports and medical records were over 99% for total breast cancer, invasive breast cancer and ER-positive breast cancer. Thus, self-reported information is used when medical records were not obtained. Currently, 2,081 incident postmenopausal cases (n=1,589 invasive) have been reported among SS participants (before analytic exclusions).

3.6 Statistical approaches

The first step in the dissertation work was to develop the ERDP using data from the nested PLCO study (Aim #1). After the ERDP was derived, it was applied in a prospective investigation with postmenopausal breast cancer risk among women randomized to the intervention arm of PLCO (Aim #1) and women enrolled in SS (Aim #2). The final application of the ERDP in the dissertation work was incorporating it into the ERLS (Aim #3). In a similar fashion to the ERDP, the ERLS was used in a prospective investigation of an association with postmenopausal breast cancer risk among women randomized to the intervention arm of PLCO (Aim #3). Development of the ERDP and ERLS, along with a description of how they were used in prospective analyses are described in the following sections. All statistical tests and models were performed in SAS 9.4 (SAS Inc., Cary, NC) using two-sided tests with α=0.05.

3.6.1 Derivation of estrogen-related dietary pattern

Unconjugated estradiol and the 2/16 ratio were identified *a priori* for inclusion because of the cumulative evidence, particularly from recent studies using the advanced

LC/MS-MS, which has supported their role in the development of postmenopausal breast cancer.[11,35,84,255] Furthermore, these two EMs were associated with postmenopausal breast cancer risk in the nested PLCO study used to derive the ERDP.[35] It is hypothesized that unconjugated estradiol the 2/16 ratio are representative of total exposure from circulating estrogens as well as the competing metabolic pathways which are suspected to have opposing influences with regards to breast cancer risk, respectively.[11] The bulk of this evidence, as well as biologic plausibility of the hypothesized relationships, have been presented in section 2.1.

To identify foods that are correlated with unconjugated E2 and the 2/16 ratio, RRR modeling was applied to the subsample of 653 participants with EM data. An approach using RRR determines linear functions of predictors, which in the present case are food groups, by maximizing the explained variation in multiple disease-specific response variables, comprised of E2 and the 2/16 ratio.[256] The primary benefit of using RRR in nutritional epidemiology is it combines data-driven and hypothesis-driven approaches into one.[31] The hypothesis-driven aspect comes from the response variables that are predefined by the investigators to be important mediators in disease risk. The data-driven aspect comes from identification of predictor variables, or food groups, which explain the greatest variation in EMs specific to our study population. Previous comparisons of RRR with other data-driven methods, such as principal component analysis, have shown stronger association with RRR in predicting cardiometabolic diseases.[31,257] A limitation of RRR is its dependence on selecting response variables that are strongly associated with disease risk. If the response variables do not mediate disease risk, it is unlikely an association between dietary factors and the disease endpoint will be

identified. However, we believed there is clear and sufficient evidence of a strong association between estrogen metabolism and breast cancer risk, as outlined in section 2.1.

In order to ensure RRR factors are based on how much variation in the outcome they explain, all intakes are centered and scaled so that their mean ± standard deviation (SD) is equal to 0 ± 1. Only the first factor was retained for development of the ERDP because it represented a dietary pattern that explains the largest variation in the EM. Initially, all 32 food groups were entered into the model at once. Those with a variable importance in projection statistic (VIP) greater than 0.8 were retained and re-entered into the RRR model, as they represented the food groups which are the strongest contributors to RRR factors scores.[258] RRR factors scores can only be calculated in participants with EM data, therefore, to apply the ERDP to the full analytic populations in PLCO and SS, we calculated the ERDP score so that it is perfectly correlated with the RRR factor scores among the subsample. To do so, food group intakes were centered and scaled, then multiplied by the corresponding model weight for each of the retained food groups, which was then summed to calculate the total ERDP score. This same calculation method was applied to score the ERDP for the full analytic cohorts in PLCO and SS. Scores with higher ERDP values theoretically represent diets with the largest collective potential to affect unconjugated E2 and the 2/16 ratio.

3.6.2 Estrogen-related lifestyle score

After the ERDP was derived and evaluated on its own, it was incorporated into the ERLS. The other lifestyle components with sufficient evidence of an effect on estrogen metabolism that completed the score were alcohol consumption, obesity status,

and PA.[12] The parameters used as criteria for scoring all of the components, with the exception of the ERDP, were similar to those outlined in the WCRF/AICR Second Expert Report, and the USDA's 2015 DGA.[13,25] Scoring criteria for the ERDP component was based on the median score for the PLCO population. Women with a score greater than or equal to the median received a 0, as those diets were hypothesized to be positively associated with estrogen metabolism and subsequent breast cancer risk. Women with an ERDP score below the median received a 1. Due to the strength of evidence for associations between alcohol intake and obesity status with breast cancer risk, and robust evidence of an estrogenic effect, they were given a stronger weight in the scoring of the ERLS by assigning women to one of three levels instead of only two levels.[13] For alcohol intake, women who abstained from drinking (0 drink/week) were scored a 2; women who consumed >0 to 7 drinks/week were scored a 1; and those who consumed >7 drinks/week were scored a 0. Women were scored a 2 if they were normal weight (BMI <25.0 kg/m^2), a 1 if overweight (BMI 25.0-29.9 kg/m^2), and 0 if obese (BMI ≥30.0 kg/m^2). For PA, women who reported >2 hours/week of vigorous PA were considered active and scored a 1, and those who reported ≤2 hours/week were scored a 0. The score for each of the four different ERLS components was summed. Women with the minimum score of 0 were hypothesized to have the largest risk profile, and those with a maximum of 6 were hypothesized to have the lowest collective risk profile from estrogen-related lifestyle factors. A summary of the ERLS scoring is portrayed in **Table 3.2**.

3.6.3 Prospective investigations

The methods used in the prospective application of the ERDP in both study populations and of the ERLS in PLCO were principally the same. The primary exposure

in Aims #1 and #2 was the ERDP, and the ERLS in Aim #3. The primary outcome for all prospective investigations was postmenopausal breast cancer followed by investigations of ER subtypes of postmenopausal breast cancer. Using descriptive statistics, study participants for the full intervention arm of PLCO and SS were characterized in terms of potential confounders and effect modifiers within strata of ERDP score quartiles in Aims #1 and #2. In Aim #3, women from the intervention arm of PLCO were characterized by categories of ERLS score (0-2; 3; 4; 5-6). Statistical comparisons of ERDP quartiles and ERLS categories were performed using t-tests and chi-square tests for continuous and categorical variables, respectively.

3.6.3a The ERDP

In time-to-event analyses, the association between breast cancer and ERDP scores were determined in Aims #1 and #2. The lowest ERDP quartile served as the referent, representing diets least associated with estrogen. Cox proportional hazards models were used to analyze the relationship between ERDP scores and incident breast cancer events, with person-time contributed as time scale variable. A test for the proportional hazards assumption was performed by inclusion of an interaction term between exposure with follow-up time, log of follow-up time, and was evaluated using Martingale-based residuals. Estimates of associations were presented as HRs with 95% CIs. An initial model was performed with adjustment for age and a second model adjusted for age and total caloric intake. The third model included multivariable adjustment. Potential confounders were selected based on a directed acyclic graph (DAG) for the hypothesized relationship between the ERDP and postmenopausal breast cancer (**Figures 3.2 & 3.3**) as well as evidence from the previous literature and model selection procedures. According

to the DAG, age, education, PA, and BMI in young adulthood represent the minimally sufficient set of confounders to include. Demographic factors of age, education, race/ethnicity and study center were included in the multivariable-adjusted models, along with total caloric intake for their putative roles as confounders for breast cancer. The remaining covariates included in multivariable-adjusted models were chosen using stepwise model selection for each of the aims with entry/exit criteria of p=0.2 to improve model efficiency and reduce the potential for over adjustment. Potential confounders for the stepwise model selection included: baseline BMI, BMI in young adulthood, HRT, OC use, family history of breast cancer, smoking status, bilateral oophorectomy, prior hysterectomy, parity, age at first birth, age at menarche, and age at menopause. Categorization of each of the potential confounders are described in more detail in the chapters corresponding to each aim (4, 5, & 6). Each of the previously described three models were performed within strata of ER subtype. A competing risk model was performed to assess a differential association for the ERDP on ER+ and ER- subtypes using a Wald test for heterogeneity in the stratified Lunn-McNeil approach.[259] A sensitivity analysis was performed to evaluate BMI as a potential mediator in the association between the ERDP and breast cancer. BMI was omitted from the full Cox proportional hazards model in order to assess the potential for mediator bias. It is possible BMI lies on the causal pathway between ERDP and breast cancer, as seen in **Figure 3.2**, therefore adjusting for it would be inappropriate and introduce bias into the association. Other sensitivity analyses related to the respective study populations were also conducted and are described further in the chapters corresponding to each analytic aim.

Lastly, the role of potential effect modifiers was assessed in the final Cox proportional hazards models. It is hypothesized that other estrogen-related risk factors may modify the association between the ERDP and breast cancer as a result of their relative estrogenic effects. To assess effect modification, an interaction term was included in the model between the ERDP and the following risk factors: BMI, HRT, alcohol consumption, parity, and PA.

Unmeasured confounding may introduce bias into epidemiologic investigations. While the PLCO and SS collected data on all known risk factors for breast cancer, it is possible that early life nutritional factors may confound the relationship between the ERDP and breast cancer. The SS has information on early life diet, but PLCO does not. Therefore, we planned to use effect estimates between early life diet in SS to estimate the potential for unmeasured confounding in PLCO using the methods proposed by VanderWeele et al.[260] However, intake of meat, plant, and fish servings at age 10 were not associated with postmenopausal breast cancer in SS, nor was a vegetarian diet before the age of 21 associated (data not shown), therefore these methods were not applied. It is possible that the recall of intake at these younger ages was afflicted by measurement error, which may partially explain the null associations and would argue for pursuing unmeasured confounding analyses. However, this type of analyses requires the estimation of effect for the unmeasured confounders which is difficult to assess from the available literature. Furthermore, both study populations contained data on BMI as a young adult, which has been shown to have an inverse association with breast cancer.[24] Therefore, we were able to account for one early life nutritional factor with a known influence on breast cancer risk.

3.6.3b The ERLS

Many of the same techniques used to investigate the ERDP in Aims #1 and #2 were used to investigate the ERLS in Aim #3, with slight differences with respect to confounders and effect modifiers. Again, a time-to-event analysis was performed with person-time contributed as time scale variable and the lowest ERLS scores of 0-2 as the referent. According to the DAG, age, BMI at young adulthood, education, and race/ethnicity are the minimally sufficient set needed for adjustment. Three different Cox proportional hazards models were performed: an initial model with adjustment for age, a second model with adjustment for age and total caloric intake, and a final model with multivariable adjustment for potential confounders. The same approach to confounders outlined above in the ERDP studies was used, with the exception of those included in the development of the ERLS (alcohol, BMI, and PA). As with the ERDP, the previously described three models were performed within strata of ER subtype in order to assess a differential effect using a competing risk model, and effect modification by other potential estrogen-related risk factors was evaluated using a stratified approach.

3.6.3c Power Calculations

Estimates for statistical power over a range of effect sizes and baseline probabilities of disease are displayed in **Table 3.3**. The NCI's Power software was utilized in all calculations, with $\alpha=0.05$.[261] The analytic study population, described in section 4.3.1 includes 27,488 participants. The analytic population in SS, as described in section 5.3.1, includes 37,752. As shown in the table, we had sufficient power to detect moderately small effect sizes (≥1.2) across a range of baseline disease probabilities. The baseline probability of breast cancer in both populations is around 6%.

3.7 Limitations and strengths

The present work was susceptible to minor, yet reconcilable limitations. As with most prospective nutritional epidemiologic studies, there was the potential for bias due to the selection of subjects, loss to follow-up, and dietary measurement error. Although FFQs may not generate accurate estimates for absolute intakes of nutrients, they have been shown to be effective in ranking individuals, as was the purpose in this study.[173] As previously mentioned, the two different FFQs used in the PLCO and SS may introduce some bias as a result of slight differences in the measurement of certain foods. There is always potential for unmeasured confounding, but the use of studies designed to investigate relationships with cancer helped to provide complete information on any known confounders.

It was possible the food pattern derived from a subsample of PLCO participants would not result in an association with breast cancer risk in the full PLCO screening arm. However, we planned to test the association in the SS to see if a lack of an association held true for another group of women. If no association was identified in both the PLCO and SS, it was possible the ERDP may still contribute to risk as a part of a lifestyle score, which was evaluated in Aim #3. It also is possible that the ERDP does not explain enough variation in EMs to influence breast cancer risk on its own, therefore we incorporated it into a lifestyle score to assess its influence with other estrogen-related risk factors. A minor limitation in regards to study populations was the lack of heterogeneity of race and ethnicities. However, our populations are predominately non-Hispanic White women, who experience the highest incidence of breast cancer so our results have public health significance.

There are many strengths in the approach and design to offset some of the limitations in the proposed research. The use of large, prospective cancer cohorts allowed for the associations of interest to be investigated with complete information on known confounders and enough power to detect moderately small effects. The application of RRR to derive the ERDP has shown larger associations than other data-driven methods in nutritional epidemiology while also incorporating hypothesized pathogenic pathways.[31,257] A major improvement upon the previous estrogen-correlated dietary pattern was in our assessment of EMs. EMs were measured using a more sensitive assay, and EMs which have been shown to be most strongly related to breast cancer risk were used in the RRR.

The resulting information from the proposed dissertation work will help to address a critical gap in translational breast cancer research. The burden of breast cancer is far-reaching, as it remains the most frequently diagnosed cancer and one of the most fatal cancers among women.[2] There is still a major need to identify primary prevention methods for breast cancer, and investigations into diet to date have been inconclusive.[43] Derivation of a dietary pattern evaluating the influence of diet as a whole based on a plausible mechanistic pathway may help to resolve the inconsistencies in previous studies. Overall, the present dissertation research contributes much-needed information about risk factors for a relatively common cancer among women, and potentially identifies novel intervention targets for primary prevention

3.8 Tables and figures

Table 3.1 Food groups used in the development of the estrogen related dietary pattern (ERDP)

Food Group	Units/day
Whole grain	ounces
Non-whole/refined grain	ounces
Dark-green vegetables	cups
Cruciferous vegetables	cups
Orange vegetables	cups
White potatoes	cups
Other starchy vegetables	cups
Tomatoes	cups
Other vegetables	cups
Citrus fruits, melons, and berries	cups
Other fruits	cups
Milk	cups
Yogurt	cups
Cheese	cups
Meat (beef, pork, veal, lamb, game)	ounces
Organ meats (meat, poultry)	ounces
Frankfurters and luncheon meats	ounces
Poultry	ounces
Fish and shellfish high in ω-3 fatty acids	ounces
Fish and shellfish low in ω-3 fatty acids	ounces
Eggs	ounces
Cooked dry beans and peas	cups
Soybean products	ounces
Nuts and seeds	ounces
Discretionary oil	grams
Discretionary solid fat	grams
Added sugars	teaspoons
Beer	drinks
Liquor	drinks
Wine	drinks
Tea	cups
Coffee	cups

ERDP: estrogen-related dietary pattern; MPED: My Pyramid Equivalents Database.

Table 3.2 Scoring parameters for estrogen-related lifestyle score (ERLS)

ERLS factor	Score	Description
ERDP	0	≥ median ERDP score
	1	< median ERDP score
Alcohol use	0	Heavy: >7 drinks/week
	1	Moderate: >0 to 7 drinks/week
	2	Abstainer: 0 drinks/week
Weight Status	0	Obese: BMI ≥30 kg/m^2
	1	Overweight: BMI 25.0-29.9 kg/m^2
	2	Normal weight: BMI <25 kg/m^2
Physical Activity (PA)	0	Inactive: ≤2 hours/week of vigorous PA
	1	Active: >2 hours/week of vigorous PA

BMI: body mass index; ERDP: estrogen related dietary pattern; ERLS: estrogen related lifestyle score; PA: physical activity

Table 3.3 Power calculations[a]

		Baseline Probability of Breast Cancer			
Study	Effect size	0.04	0.05	0.06	0.07
PLCO	1.1	0.25	0.30	0.35	0.39
	1.2	0.72	0.80	0.86	0.91
	1.3	0.99	0.99	>0.99	>0.99
SSS	1.1	0.33	0.39	0.45	0.41
	1.2	0.84	0.91	0.95	0.97
	1.3	0.98	>0.99	>0.99	>0.99

ERDP: estrogen-related dietary pattern; ERLS: estrogen-related lifestyle score; PLCO: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; SS: The Sister Study.
[a]α=0.05; PLCO n=27,488; SS=37,752;

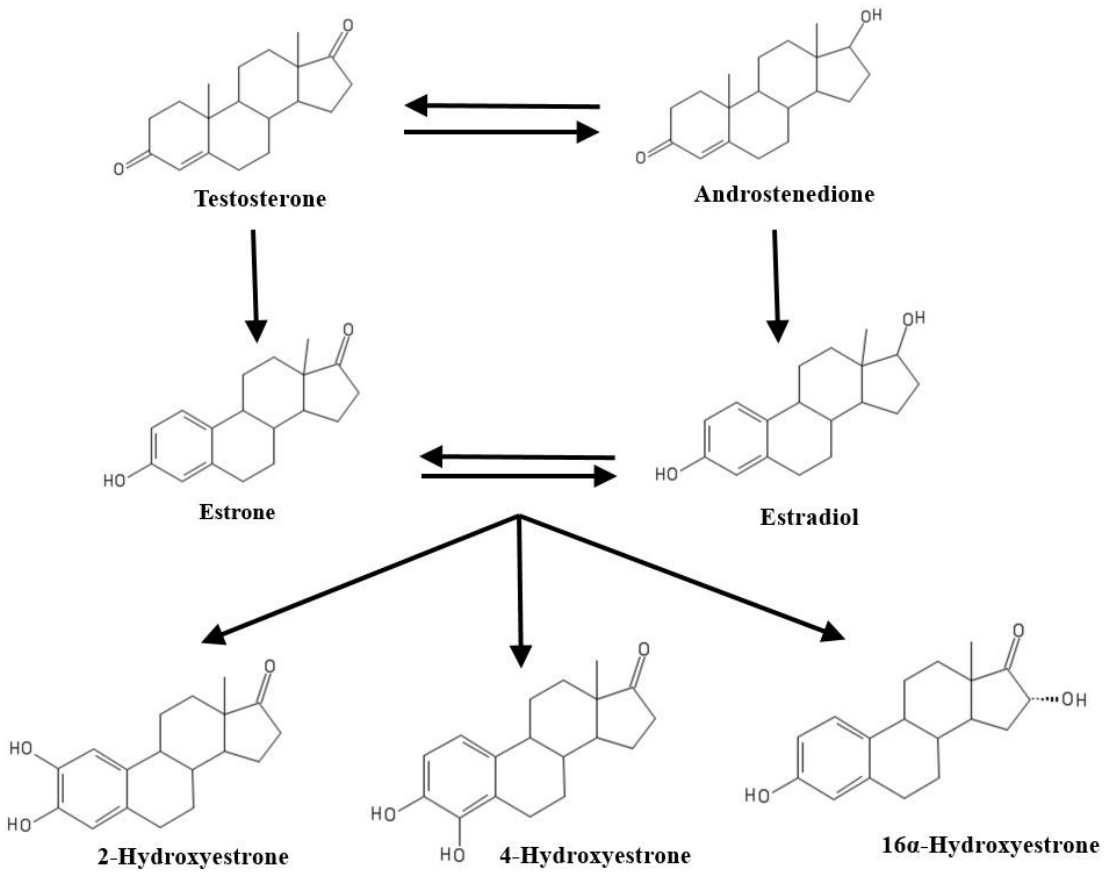

Figure 3.1 Estrogen synthesis and metabolism[a]

[a]Adapted from Furhman et al.[35]

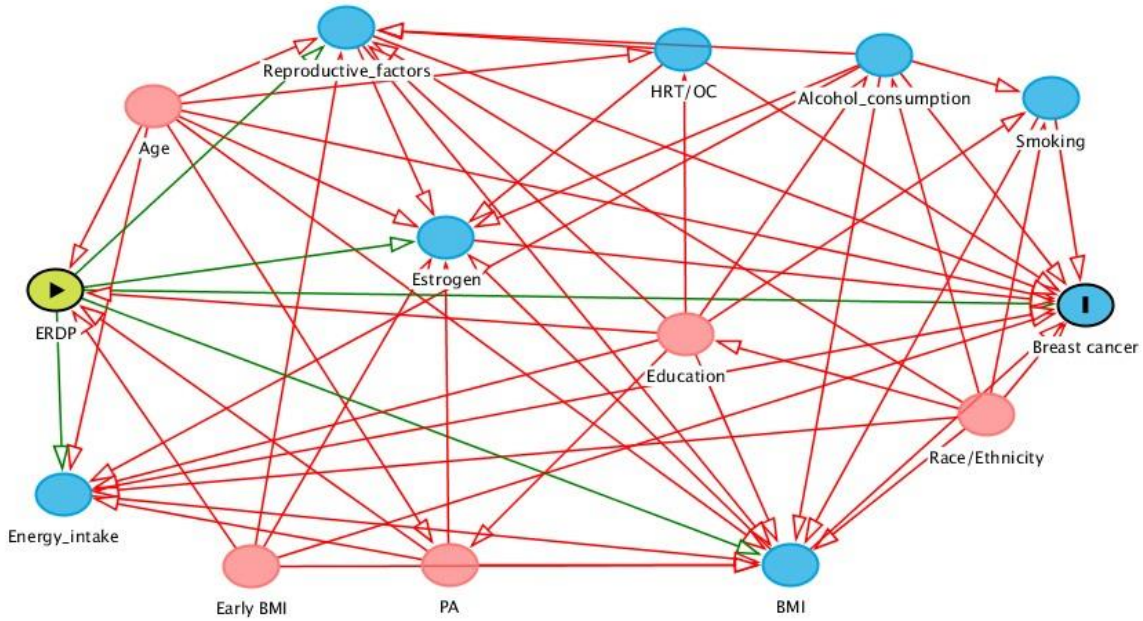

Figure 3.2 Directed acyclic graph for the association between ERDP and postmenopausal breast cancer[a]

BMI: body mass index; ERDP: estrogen-related dietary pattern; HRT: hormone replacement therapy; OC: oral contraceptive; PA: Physical activity.
[a]Parity, age at menarche, age at menopause, age at first birth, and oophorectomy/hysterectomy are included in reproductive factors.

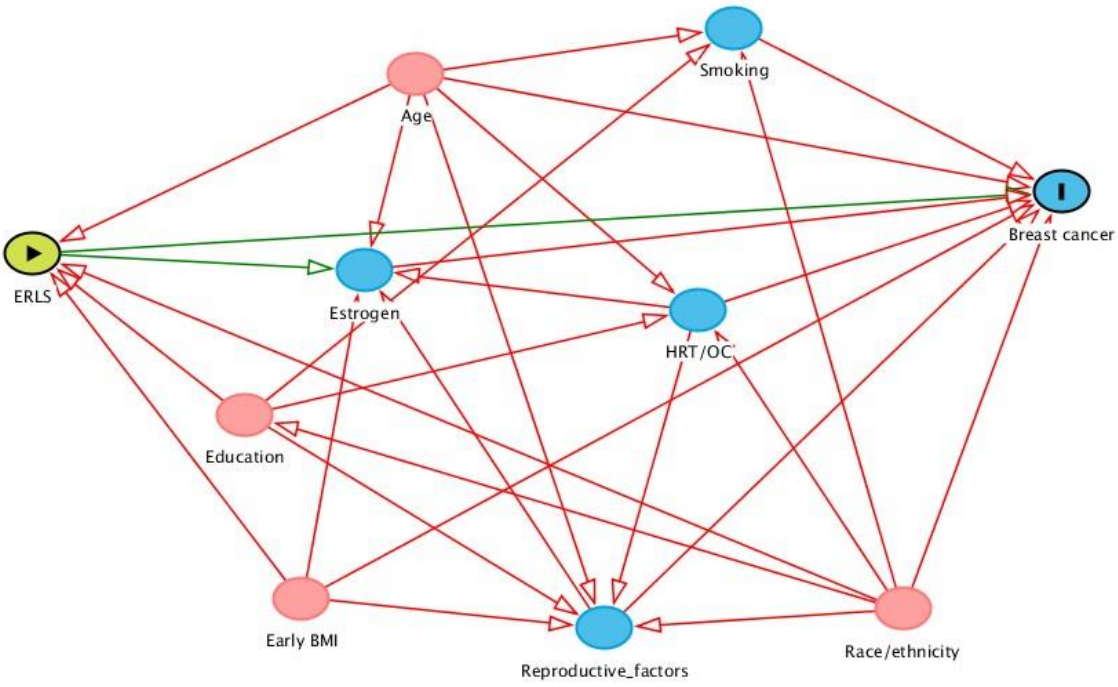

Figure 3.3 Directed acyclic graph for the association between ERLS and postmenopausal breast cancer[a]

BMI: body mass index; ERDP: estrogen-related dietary pattern; ERLS: estrogen-related lifestyle score; HRT: hormone replacement therapy; OC: oral contraceptive; PA: Physical activity.
[a]Parity, age at menarche, age at menopause, age at first birth, and oophorectomy/hysterectomy are included in reproductive factors.

CHAPTER 4

A DIETARY PATTERN BASED ON ESTROGEN METABOLISM IS ASSOCIATED WITH BREAST CANCER RISK IN A PROSPECTIVE COHORT OF POSTMENOPAUSAL WOMEN.

Mark A. Guinter,[1,2] Alexander C. McLain,[1] Anwar T. Merchant,[1] Dale P. Sandler,[3] Susan E. Steck[1,2]

[1]Department of Epidemiology and Biostatistics; [2]Cancer Prevention and Control Program, Arnold School of Public Health, University of South Carolina, Columbia, SC; [3]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC.

Submitted to *International Journal of Cancer*

4.1 Abstract

Increased exposure to estrogen is an established risk factor for postmenopausal breast cancer, and dietary factors have been shown to influence estrogen metabolism. However, investigations of diet and breast cancer have been inconclusive. We developed a dietary pattern associated with levels of unconjugated estradiol and the ratio of 2- and 16-hydroxylated estrogen metabolites in a subsample of Prostate, Lung, Colorectal and Ovarian Screening Trial (PLCO) participants (n=653) using reduced rank regression, and examined its association with postmenopausal breast cancer prospectively in the larger PLCO cohort (n=27,488). The newly developed estrogen-related dietary pattern (ERDP) was comprised of foods with positively weighted intakes (non-whole/refined grains, tomatoes, cruciferous vegetables, cheese, fish/shellfish high in ω-3 fatty acids, franks/luncheon meats) and foods with negatively weighted intakes (nuts and seeds, other vegetables, fish/shellfish low in ω-3 fatty acids, yogurt, coffee). A 1-unit increase in the ERDP score was associated with a 9%, 13%, and 13% increase in total breast cancer risk (HR: 1.09, 95%CI: 1.01-1.18), invasive (HR: 1.13; 95%CI: 1.04=1.04-1.24) and estrogen receptor (ER)-positive (HR: 1.13, 95%CI: 1.02- 1.24) breast cancer, respectively, after adjustment for confounders. Associations were seen for the fourth quartile of ERDP for overall breast cancer (HR: 1.14; 95% CI: 0.98, 1.32), invasive cases (HR: 1.20, 95%CI: 1.02, 1.42) and ER -positive cases (HR: 1.19; 95%CI: 0.99-1.41) compared to the first. The increased risk associated with increasing ERDP score was more apparent in strata of some effect modifiers (non-hormone replacement therapy users and non-obese participants) where the relative estrogen exposure due to that factor was lowest. Our results suggest a dietary pattern based on EM is positively associated with postmenopausal breast cancer risk, possibly through an estrogenic influence.

4.2 Introduction

Breast cancer, the most commonly diagnosed cancer among women worldwide, is a disease of strong hormonal influence.[1] Serum and urinary levels of estrogen metabolites (EMs) have consistently been associated with postmenopausal breast cancer risk in prospective studies.[11] Therefore, modifiable lifestyle risk factors for postmenopausal breast cancer that are associated with estrogen metabolism may present opportunities for primary prevention.

Diet is commonly studied as a point of intervention for reducing cancer risk, however there have been conflicting results in dietary investigations into breast cancer risk, with the exception of alcohol which is considered an established risk factor.[17–19,24] It is likely that the practice of studying dietary components in isolation may contribute to the inconclusive findings for associations with breast cancer, as it does not take into account the interactions between nutrients and phytochemicals.[25] Therefore, it is beneficial to study diet in its entirety using dietary pattern analyses when investigating a potential association with breast cancer.[26] Emerging evidence has supported an association between some dietary patterns and incident breast cancer risk.[17,18,27] Many of the diets that have indicated an inverse relationship with breast cancer are characterized by high intakes of fruits and vegetables, and diets with increased risk typically have higher intakes of fat and animal products.[17,21,28]

In order to address some of the inconclusive findings in the literature on diet and breast cancer, it may be advantageous to consider the mechanistic pathway by which a potential association may occur. Nutritional factors can influence many hormonal processes in women, such as the development of breasts, and the onset of both menarche

and menopause.[29,30] Therefore, diet may have a role in altering estrogen metabolism and subsequently breast cancer risk, although data on the relationship between diet and estrogen metabolism is scarce.[13] A relatively new approach to dietary pattern analyses, reduced rank regression (RRR), allows for the use of biomarkers, such as EMs, in developing a dietary pattern that can then be investigated in association with disease endpoints.[31] Previously, Fung et al. developed a dietary pattern correlated with serum levels of estradiol and estrone sulfate using RRR, but the pattern subsequently was not associated with breast cancer among postmenopausal women in the Nurses' Health Study (NHS).[32] However, application of the same estrogen-correlated dietary pattern in a Swedish cohort identified a positive association with incident breast cancer.[27]

In the present analysis, we used RRR to develop a dietary pattern that is associated with EMs that are hypothesized to be associated with breast cancer risk. Using a liquid chromatography-tandem mass spectrometry assay (LC/MS-MS), 15 EMs can be measured in an accurate and reproducible method with enough sensitivity to detect the low levels present in postmenopausal women.[262] Measurement of the parent estrogens' downstream EMs allows for ratios of competing metabolic pathways to be quantified. There is evidence that 2-hydroxylation of the parent estrogens is inversely associated, and 16-hydroxylation is positively associated with postmenopausal breast cancer.[11] Therefore, increases in the ratio of 2- to 16-hydroxylated EMs (2/16) is hypothesized to indicate a beneficial shift in estrogen metabolism with respect to breast cancer risk.[11] Based on this evidence, and established evidence linking unconjugated estradiol (E2) to postmenopausal breast cancer risk,[11,57] we used RRR to develop a dietary pattern associated with 2/16 and E2. This newly developed estrogen related dietary pattern

(ERDP) was applied in a prospective cohort of women to examine an association with total postmenopausal breast cancer and by estrogen-receptor (ER) subtype. The potential for effect modification by other estrogen-related risk factors was examined.

4.3 Methods

4.3.1 Study Population

The Prostate, Lung, Colorectal & Ovarian Cancer Screening Trial (PLCO) is a large population-based trial designed to determine the effects of screening on cancer prognosis and mortality. Design and implementation has been described in detail elsewhere.[33] Briefly, 76,685 men and 78,216 women aged 55 to 74 were recruited at 10 different screening centers across the United States between 1993 and 2001. Eligible participants underwent a physical examination and filled out a questionnaire with information on demographics, medical history, family history, lifestyle factors, and recent history of participation in screening examinations at baseline. Follow-up continued for 13 years or until December 31, 2009. For the current study, the analysis was restricted to screening arm participants (n=39,104) as this group provided blood samples used for assessing estrogen metabolites and were asked to complete the dietary instrument (DQX). Over 82% of participants in the screening arm completed the DQX. The population was further limited to women who completed the baseline questionnaire, a valid DQX (caloric intake between 1st and 99th percentiles, <8 missing line items), and without a personal history of cancer (n=28,438). Participants were further excluded if they had an extreme body mass index (BMI) (<15 or >55 kg/m^2; n=74), if they did not contribute any

person-time (n=58) or were missing covariate data (n=818), bringing the final analytic sample to 27,488.

4.3.2 Subsample and EM Assay

A subset of postmenopausal women randomized to the screening arm of PLCO for whom information on serum EMs was available was utilized to derive the ERDP. Complete information on the nested study has been published elsewhere.[35] Briefly, the nested study population was drawn from all 1,141 incident breast cancer cases diagnosed from the start of recruitment in 1993 through June 30, 2005, and a random sample of 1,141 control subjects. After excluding women who were not postmenopausal, were using hormone replacement therapy (HRT) at baseline, or had prior diagnoses of cancer, the sample was reduced to 390 cases and 453 controls. For the purposes of the present analysis, cases who were diagnosed <2 years after serum sample donation (n=98) were excluded to avoid the possibility of disease processes affecting estrogen levels. Women without a valid DQX (n=77) or with implausible EM levels (i.e., if they were outside of 25^{th} and 75^{th} percentile, plus/minus three times interquartile range; n=15) were further excluded. The final analytic sample for the RRR procedure included 393 controls and 260 subsequent cases, with a mean of 5.25 years from sample donation to breast cancer diagnosis.

Serum samples from women in the subsample were collected at baseline, stored at −80°C and were thawed at 4°C. The LC/MS-MS assay was used to measure the parent estrogens along with their metabolites in the 2-, 4-, and 16-hydroxylation pathways, for a total of15 EMs. The specifics of sample preparation and LC/MS-MS methods have been

described elsewhere.[41] The coefficient of variation for all EMs was <5%, with even lower coefficients evident for the parent estrogens (<3%) and E2 (<2%).[35]

4.3.2 Dietary Measurement

The DQX, a 137-item food frequency questionnaire, was designed specifically for PLCO and asked about typical frequency of intake over the past year. Typical portion size was assessed for 77 of the items. Nutrient and food intake amounts were calculated using US dietary data and the pyramid food group servings database from the US Department of Agriculture (USDA).[263] Food and nutrient values were used to create food groups based on the USDA's My Pyramid Equivalents Database (MPED), with additional groups created for cruciferous vegetables, tea, and coffee.[251] The 32 groups used in the present analysis are shown in **Table 3.1**.

4.3.4 Breast Cancer Ascertainment

Incident breast cancer cases were identified primarily through self-report via annually mailed follow-up questionnaires. Other sources of ascertainment included the National Death Index, physician reports, state cancer registries, and next of kin reports. Over 96% of the cases were confirmed through hospital records.[252] In the analytic cohort, a total of 1,569 incident breast cancer cases occurred. A supplemental form was implemented in 2007 to capture more detailed information about the diagnosis, available for 98% of cases.

4.3.5 Development of the ERDP

To identify foods that are correlated with unconjugated E2 and the 2/16 ratio, RRR modeling was applied to the subsample of 653 participants with EM data. An approach using RRR determines linear functions of predictors, which in the present case are food groups, by maximizing the explained variation in multiple disease-specific response variables, comprised of E2 and the 2/16 ratio.[256] In order to ensure RRR factors are based on how much variation in the outcome they explain, all intakes were centered and scaled so that their mean ± standard deviation (SD) is equal to 0 ± 1. Only the first factor was retained for development of the ERDP because it represented a dietary pattern that explained the largest variation in the EM. Initially, all 32 food groups were entered into the model at once. Those with a variable importance in projection statistic (VIP) greater than 0.8 were retained and re-entered into the RRR model, as they represent the food groups which are the strongest contributors to RRR factors scores.[258] The model weights were extracted from the final RRR model from PROC PLS and SAS version 9.4 (SAS Inc., Cary, NC). To calculate the ERDP score in the full analytic PLCO cohort food group intakes were centered and scaled, then multiplied by their corresponding model weights (**Table 4.1**) for each of the retained food groups. The total ERDP score was calculated by summing over the weighted intakes. This same calculation method was applied to score the ERDP for the full analytic cohort.

4.3.6 Statistical Analysis

Baseline comparisons of participant characteristics by ERDP quartiles were performed using t-tests and chi-square tests for continuous and categorical variables,

respectively. Cox proportional hazards models were applied to prospectively analyze the relationship between ERDP scores and incident breast cancer events, with person-time contributed as a time scale variable. ERDP scores were categorized into quartiles, with the first quartile set as the referent. The first quartile hypothetically represents diets with an estrogen profile associated with the lowest breast cancer risk (low levels of unconjugated E2 and high 2/16 ratio). The hazard ratio and 95%CI also were calculated for the continuous ERDP score variable, and the p-value reported as a test for trend. Covariates for multivariable adjusted models were chosen using stepwise model selection with entry/exit criteria of p=0.2. We adjusted for age (years), HRT use (current; former; never), body mass index (BMI) (kg/m^2), alcohol consumption (abstainer; 1-7; >7 drinks/week), family history of breast cancer (yes; no), education (less than high school; high school and some college; college degree; graduate degree), bilateral oophorectomy (yes; no), parity (6 categories), age at menopause (5 categories), hours of vigorous physical activity per week (6 categories), and total energy intake (kcal/day). Age at first birth, age at menarche, oral contraceptive use, race/ethnicity, smoking status, and prior hysterectomy were also considered as potential confounders but were not included after performing the stepwise model selection. The potential for effect modification by BMI (18.5-29.9 kg/m^2; ≥30 kg/m^2), baseline HRT use (yes; no), alcohol consumption (<1 drink/week; ≥1/week), parity (nulliparous; parous), and vigorous physical activity per week (<2 hours; ≥2 hours) was assessed using a multiplicative interaction term in the model. All models were performed with overall breast cancer and by ER subtype. A competing risk model was used to assess a differential association for the ERDP on ER+

and ER- subtypes using a Wald test for heterogeneity in the stratified Lunn-McNeil approach.[259]

4.4 Results

Unconjugated E2 and the 2/16 ratio were moderately and inversely correlated (r= -0.51; p<0.0001) in the subsample of 653 women. After applying the VIP criteria, 11 food groups with a VIP >0.8 were retained and re-entered into the RRR procedure. The final list of food groups included in the ERDP is shown in **Table 4.1.** Overall, 4.9% of the variation in the EMs was explained by the ERDP. Intakes of non-whole/refined grains, tomatoes, cruciferous vegetables, cheese, fish/shellfish high in ω-3 fatty acids, and franks/luncheon meats were added; and intakes of nuts and seeds, other vegetables, fish/shellfish low in ω-3 fatty acids, yogurt, and coffee were subtracted to calculate the ERDP score. The "other vegetables" group includes vegetables except for tomatoes, potatoes and orange, dark leafy, cruciferous, and starchy vegetables. For example, this group includes cucumber, onion, green pepper, beet, celery, and lettuce. The resulting ERDP scores were weakly but significantly correlated with unconjugated E2 (r=0.27; p<0.0001) and the 2/16 ratio (r=-0.16; p<0.0001) (**Table 4.2**). When considering the intakes of ERDP food groups, the strongest correlates with unconjugated E2 were non-whole/refined grains (r=.10; p=0.01), cheese (r=0.16; p<0.0001), yogurt (r=-0.10; p=0.01), and franks/luncheon meats (r=0.11; p=0.001). Only intakes for non-whole/refined grains (r=-0.09; p=0.02) and cheese (r=-0.08; p=0.05) were significantly correlated with the 2/16 ratio. Increasing ERDP scores are positively correlated with unconjugated E2 and negatively correlated with the 2/16 ratio.

Table 4.1 compares the mean intakes of included food groups across extreme quartiles of unconjugated E2 and the 2/16 ratio. On average, participants in the highest quartile of unconjugated E2 consumed higher amounts of non-whole/refined grains (4.45 vs. 3.90; p=0.01), cheese (0.43 vs. 0.29; p<0.01), and franks/luncheon meats (0.34 vs. 0.21; p=0.01) compared to participants in the first quartile. Mean consumption of coffee (2.30 vs. 3.09; p=0.04) and yogurt (0.08 vs 0.12; p=0.03) were significantly lower among participants in the highest quartile of unconjugated E2 compared to the first. There were no significant differences in mean intakes when comparing extreme quartiles of the 2/16 ratio.

There were 1,592 confirmed incident cases of breast cancer (n=1,248 invasive) over an average follow-up of 10.9 years. Among the cases, 1,097 were ER+ and 189 were ER-. The mean ± SD ERDP score was -0.006± 0.646 with a range of -4.515 to 6.578. Women who were diagnosed with breast cancer during follow-up had significantly higher mean ERDP scores at baseline compared to women who were not diagnosed during follow-up (0.037 vs. -0.009, respectively; p=0.006). Baseline characteristics for the full analytic cohort, stratified by ERDP quartiles, are shown in **Table 4.3**. There was a stepwise increase in the number of total cases from the first to fourth quartiles although the differences across quartiles was not significant (p=0.12). Women in the fourth quartile of the ERDP were younger, had a higher mean BMI, higher daily caloric intake, were more likely to have had a bilateral oophorectomy, and were more likely to be non-Hispanic White compared to women in the first quartile. There was no clear trend for alcohol, with a higher proportion of both abstainers and heavier drinkers in the highest quartile of ERDP. A similar pattern was seen for physical activity. There were no

differences in HRT use, parity, family history of breast cancer, or age at menopause across ERDP quartiles. Participants in the highest quartile of ERDP score consumed the most non-whole/refined grains, tomatoes, cheese, and franks/luncheon meats. On the contrary, participants in the lowest quartile consumed the most coffee, nuts and seeds, fish/shellfish low in ω-3 fatty acids, yogurt, and other vegetables.

Results from the time-to-event analyses are shown in **Table 4.4**. In models using ERDP quartiles, participants in the fourth quartile were at increased risk of postmenopausal total breast cancer (HR: 1.14; 95% CI: 0.98, 1.32) and invasive breast cancer (HR: 1.20; 95%CI: 1.02, 1.42) after multivariable adjustment. All quartiles were positively associated with risk, with increasing magnitude of effect estimates with increasing quartiles, compared to the first for total (p-trend=0.04) and invasive breast cancer (p-trend=0.005). The continuous ERDP variable was positively associated with total and invasive breast cancer risk. A 1-unit increase in ERDP was associated with a 9% increase in risk (HR: 1.09; 95% CI: 1.01, 1.18) for total and 13% increase in risk for invasive (HR: 1.13; 95%CI: 1.04, 1.24) after multivariable adjustment.

The ERDP was associated with ER+ but not ER- breast cancer (**Table 4.4**). The multivariable effect estimates for continuous ERDP were 1.13 (95%CI: 1.02-1.24; p-trend=0.02) and 1.07 (95%CI: 0.85-1.35; p-trend=0.54), respectively. The competing risk model did not indicate evidence of a differential effect of the ERDP by ER subtypes (p=0.87; data not shown).

There was no evidence for effect modification by alcohol consumption and PA. However, there was some indication that HRT, BMI, and parity may modify the effect of the ERDP (**Table 4.5**). In stratified results, estimates of association were higher in strata

of some effect modifiers where estrogen exposure is thought to be lowest (e.g., among HRT non-users, and participants with lower BMI). In the case of parity, estimates were higher in nulliparous women.

4.5 Discussion

We developed a dietary pattern that was significantly associated with serum levels of unconjugated E2 and the 2/16 ratio in postmenopausal women. Intakes of non-whole/refined grains, cheese, franks/luncheon meats, and yogurt were most strongly correlated with the derived pattern. When applied in a prospective cohort of women, the ERDP was positively associated with total and invasive postmenopausal breast cancer risk, and the association was present in ER+ but not ER- breast cancer. The risk associated with high ERDP scores was higher within strata of some effect modifiers hypothesized to have lower exposure to estrogen. These results suggest that women who consume a diet that adheres to higher ERDP scores may be at moderately increased risk of developing postmenopausal breast cancer, possibly through an influence on estrogen metabolism.

This is the first study to develop a dietary pattern based on estrogen metabolism that is specific to breast cancer risk, due to inclusion of the 2/16 ratio. Quantification of estrogen's downstream metabolic pathways that may be indicative of breast cancer risk was possible through use of a highly sensitive LC/MS-MS assay. Previously, Fung et al. used RRR to derive a dietary pattern correlated to estradiol and estrone sulfate. High scores for the pattern were characterized by high intakes of red meat, legumes, and pizza; and low intakes of whole grains and coffee. In the MPED food groups used in the ERDP,

food items that make up mixed dishes are decomposed into their individual food groups, (for example, pizza is decomposed into cheese, tomatoes, and refined grains). We observed moderate similarities between the ERDP and Fung et al.'s estrogen pattern with regard to cheese and tomatoes (in the form of pizza in Fung et al.'s pattern), coffee, and their respective directions of association with the derived patterns. Fung et al. observed an inverse association between whole grains and estrogen, and although whole grains were not a significant contributor to the ERDP, non-whole/refined grains had a significant positive association, suggesting the importance of choosing whole grains and limiting processed grains.

Other literature on dietary patterns and estrogen metabolism is scarce. However, the Alternate Healthy Eating Index and the Western pattern, comprised of processed foods and animal products, have been inversely and positively associated with estradiol, respectively.[144] An intervention study using the Mediterranean Diet, usually high in fruits and vegetables, legumes, oils, and other foods that result in a higher proportion of unsaturated fats compared to saturated fats, reported a roughly 40% decrease in total urinary estrogen levels ($p<0.02$) in postmenopausal women, showing some anti-estrogenic properties.[147] Although there is evidence linking alcohol[52] and soy products[130] with estrogen metabolism, they were not included in the ERDP because they failed to meet the inclusion criteria of a VIP >0.8 in the first RRR model. This indicated these groups did not explain a large enough variation in the EMs, possibly due to a small range of intakes for these groups in our subsample of women.

Evidence of a moderate but significant association between the ERDP and postmenopausal breast cancer was observed in our study population. A significant

association was limited to ER+ subtypes, possibly due to an influence on estrogen metabolism. Fung et al.'s estrogen diet pattern was not associated with total postmenopausal breast or ER subtype-specific cancer risk in NHS,[32] which the authors concluded was a result of the low correlation between their pattern and the estrogens (r=0.22 and r=0.24 for estradiol and estrone sulfate, respectively), which may be insufficient to affect breast cancer risk. However, when the same pattern was applied in the Swedish Mammography Cohort (SMC) a 29% increase in risk of developing breast cancer (HR: 1.29; 95% CI: 1.08, 1.55) was observed when comparing women in the highest quartile with the lowest, and no heterogeneity was observed between the ER subtypes.[27] The authors cited a wider range of intakes, higher consumption of coffee, and lower levels of other breast cancer risk factors in SMC as reasons for results that differed from the NHS. Our results are consistent with those of the SMC. Explanations for different results between the previous studies and ours are difficult to discern because of our use of different EMs which resulted in a different dietary pattern. The use of LC/MS-MS to accurately quantify the EMs, and inclusion of the 2/16 ratio that has more consistently been associated with breast cancer risk than other EMs is a strength of our investigation.

Qualitative evidence of effect modification by HRT, BMI, and parity in the association between the ERDP and postmenopausal breast cancer risk was observed. Based on prior evidence, we expect women who are not using HRT or who are not obese to have lower lifetime exposure to estrogen.[67] In these women, a dietary influence, through estrogen or other pathway may be easier to detect than in women with higher lifetime estrogen exposure. In the NHS, no effect modification by BMI was observed

using their estrogen correlated dietary pattern, though other effect modifiers were not examined.[32] It is possible a woman's nulliparity is a result of low fertility due to low hormone levels.[264] However, nulliparous women typically experience more menstrual cycles, resulting in greater lifetime exposure to estrogen and higher breast cancer risk,[219] therefore these results need to be explored further.

There are multiple possible mechanisms by which the ERDP effects estrogen metabolism and breast cancer risk, such as through influences on microbiome diversity.[265] The intestinal microbiome is strongly influenced by dietary behaviors, and the composition of the microbiome can have implications on many important physiological processes.[266] The fate of conjugated, or inactive, estrogens is dependent on the state of the intestinal microbiome, which influences whether or not the conjugated estrogens are excreted through feces or transformed to their unconjugated forms and subsequently reabsorbed.[267] If reabsorbed, there is a greater estrogenic exposure throughout the body. Therefore, diet may influence development of a microbiome that is favorable to excretion of estrogens, lowering breast cancer risk, or one that is conducive to reabsorption of the estrogens which increases risk. In addition to absolute exposure to estrogen, the composition of EMs is also influenced by the micriobiome.[268] More specifically, there is evidence of microbial effects on interconversions of the parent estrogens and hydroxylation down the 16-pathway from in vitro and human studies.[269,270] The intestinal microbiome is strongly influenced by fiber intake, or lack thereof, through consumption of grains and vegetables, both of which are included in the ERDP.[266] The ERDP also is comprised of animal products, such as meats, cheese, and yogurt, which can impact microbiome diversity.[271–273] Considering the presence of a microbial influence

on estrogen metabolism and its established relationship with diet, modification of the intestinal microbiome is a plausible mechanism by which the ERDP influences estrogen metabolism and breast cancer risk.

Considering other mechanisms, it is possible the ERDP was associated with breast cancer through effects on inflammation. Coffee, as well as processed meats, dairy, and refined grains which are common in the Western diet, have all exhibited associations with inflammation,[274,275] and inflammation may play a role in mammary tumor development.[276] The Mediterranean Diet, characterized by foods with anti-inflammatory properties has been inversely associated with breast cancer,[48] and a dietary pattern based on inflammatory potential has shown evidence of an association with breast cancer[50] and breast cancer mortality.[51]

There are some limitations in our study that need to be considered when interpreting the results. As with most prospective nutritional investigations, there is the potential for bias due to the selection of subjects, loss to follow-up, and dietary measurement error. Although food frequency questionnaires may not generate accurate estimates for absolute intakes of nutrients, they have been shown to be effective in ranking individuals, as is the purpose in this study.[173] Unexpected results from fish with low and high ω-3 fatty acids could have been due to preparation methods that were not ascertained. Low numbers of ER- cases may have limited our ability to detect an association in this subtype and a heterogeneity in effect by ER subtype, however, there were ample ER+ cases for analyses. A limitation of the PLCO study population is the lack of racial/ethnic diversity. However, non-Hispanic White women experience the

highest incidence of breast cancer compared to other races/ethnicities in the US, so results are generalizable to this group at the highest risk.

There are strengths in the approach and design to note, as well. The use of a large, prospective cancer cohort allowed the associations of interest to be investigated with enough power to detect moderately small effects and with information on multiple known risk factors with which to adjust for potential confounding. The application of RRR to derive the ERDP provides the ability to incorporate a hypothesized pathogenic pathway in dietary pattern development.[31,257] As noted, the EMs included in the RRR models have been shown to be strongly related to breast cancer risk and were measured using a more sensitive assay method, thus improving upon the previous RRR-derived estrogen dietary pattern.[32]

In conclusion, we identified a dietary pattern to be associated with an estrogen profile (high E2 and low 2/16 ratio) hypothesized to increase breast cancer risk. Women who had high ERDP scores tended to consume higher amounts of non-whole/refined grains, tomatoes, cheese, franks/luncheon meats; and lower amounts of nuts and seeds, cruciferous vegetables, other vegetables, fish/shellfish, yogurt, and coffee. A subsequent prospective investigation indicated that this estrogenic diet was associated with an increased risk of postmenopausal breast cancer risk, possibly through an influence on estrogen metabolism. Future studies should be conducted in populations from other regions with larger variation in intakes in food groups, or in study populations using open-ended dietary assessment tools to capture all foods or food groups that potentially influence estrogen metabolism

4.6 Tables

Table 4.1 Comparison of mean (±standard deviation) food or beverage intake across extreme quartiles of estrogen metabolites for the eleven foods and beverages included in the estrogen-related dietary pattern (ERDP)

		Unconjugated E2			2/16 Ratio		
	Model Weight[a]	Q1 (n=164)	Q4 (n=163)	p-value[b]	Q1 (n=163)	Q4 (n=163)	p-value[b]
Non-whole/refined grains (oz/day)	0.12	3.90 ± 1.81	4.45 ± 2.00	0.01	4.39 ± 2.06	4.10 ± 1.91	0.18
Tomatoes (cups/day)	0.09	0.40 ± 0.22	0.45 ± 0.27	0.06	0.43 ± 0.30	0.43 ± 0.24	0.79
Other vegetables (cups/day)	-0.13	0.96 ± 0.52	0.95 ± 0.45	0.89	1.02 ± 0.58	1.07 ± 0.60	0.42
Cruciferous vegetables (cups/day)	0.08	0.28 ± 0.21	0.26 ± 0.20	0.6	0.30 ± 0.26	0.32 ± 0.23	0.46
Cheese (cups/day)	0.16	0.29 ± 0.23	0.43 ± 0.38	<0.01	0.38 ± 0.37	0.33 ± 0.26	0.2
Yogurt (cups/day)	-0.12	0.12 ± 0.19	0.08 ± 0.15	0.03	0.09 ± 0.17	0.12 ± 0.21	0.15
Fish/shellfish high in ω-3 fatty acids (oz/day)	0.2	0.16 ± 0.17	0.15 ± 0.15	0.55	0.15 ± 0.17	0.16 ± 0.19	0.53
Fish/shellfish low in ω-3 fatty acids (oz/day)	-0.27	0.53 ± 0.47	0.49 ± 0.38	0.46	0.46 ± 0.36	0.52 ± 0.49	0.21
Franks and luncheon meats (oz/day)	0.08	0.21 ± 0.23	0.34 ± 0.56	0.01	0.28 ± 0.31	0.22 ± 0.28	0.07
Nuts and seeds (oz/day)	-0.11	0.45 ± 0.70	0.38 ± 0.42	0.32	0.44 ± 0.79	0.44 ± 0.64	0.99
Coffee (cups/day)	-0.1	3.09 ± 3.59	2.30 ± 3.27	0.04	2.64 ± 3.34	3.18 ± 3.73	0.17

ERDP: estrogen related dietary pattern; EM: estrogen metabolite
[a]Model weight from final RRR model that is used for ERDP scoring.
[b]t-test for the comparison of means in the first and fourth quartiles.

Table 4.2 Correlations for the estrogen-related dietary pattern (ERDP) and food group intakes with factor score and estrogen metabolite response variables among subset of women with estrogen metabolite values (n=653)[a]

	RRR Factor Score	Unconjugated Estradiol	2/16 Pathway Ratio
Total ERDP score	1.00 (<0.01)	0.27 (<0.01)	-0.16 (<0.01)
Non-whole/refined grains (oz/day)	0.41 (<0.01)	0.10 (0.01)	-0.09 (0.02)
Tomatoes (cups/day)	0.28 (<0.01)	0.08 (0.03)	-0.03 (0.48)
Other vegetables (cups/day)	-0.07 (0.07)	-0.03 (0.46)	0.00 (0.90)
Cruciferous vegetables (cups/day)	-0.02 (0.61)	-0.03 (0.47)	-0.04 (0.34)
Cheese (cups/day)	0.55 (<0.01)	0.16 (<0.01)	-0.08 (0.05)
Yogurt (cups/day)	-0.34 (<0.01)	-0.10 (0.01)	0.05 (0.25)
Fish/shellfish high in ω-3 fatty acids (oz/day)	-0.06 (0.10)	-0.02 (0.56)	0.00 (0.98)
Fish/shellfish low in ω-3 fatty acids (oz/day)	-0.18 (<0.01)	-0.04 (0.27)	0.04 (0.32)
Franks and luncheon meats (oz/day)	0.39 (<0.01)	0.11 (<0.01)	0.00 (0.92)
Nuts and seeds (oz/day)	-0.14 (<0.01)	-0.05 (0.17)	-0.05 (0.24)
Coffee (cups/day)	-0.22 (<0.01)	-0.06 (0.10)	0.03 (0.51)

ERDP: estrogen related dietary pattern; EM: estrogen metabolite; RRR: reduced rank regression
[a]Pearson's correlation coefficient (p-value).

Table 4.3 PLCO population characteristics across the estrogen-related dietary pattern (ERDP) quartiles

		ERDP Quartile (score range)			
		1st (-4.515, -0.350)	2nd (-0.351, -0.021)	3rd (-0.022, 0.328)	4th (0.329, 6.578)
n		6,872	6,872	6,872	6,872
Breast cancer cases	Total	366	392	403	431
	Invasive	280	309	331	348
	ER+	246	275	274	302
	ER-	45	41	55	48
ERDP score (mean ± SD)		-0.77 ± 0.43	-0.18 ± 0.09	0.14 ± 0.10	0.77 ± 0.48
Age (mean ± SD)		62.6 ± 5.3	62.8 ± 5.4	62.5 ±5.3	61.8 ± 5.2
BMI (kg/m^2; mean ± SD)		26.6 ± 5.1	26.6 ± 5.1	27.1 ± 5.3	28.1 ± 5.9
BMI at age 20 (kg/m^2; mean ± SD)		21.4 ± 2.9	21.1 ± 2.7	21.2 ± 2.7	21.4 ± 3.0
Total energy intake (kcal/day; mean ± SD)		1,691 ± 578	1,542 ± 528	1,659 ± 535	2,078 ± 621
HRT use (%)					
	Current	51.7	51.9	51.6	51.6
	Former	15.8	16.4	16	15.8
	Never	32	31.3	32	32.1
Race (%)					
	White, Non-Hispanic	88.4	90.6	91.9	93.1
	Black, Non-Hispanic	4.8	4.5	4.1	3.9
	Hispanic	1.1	1.2	1.4	1.4
	Asian	5.1	3.1	2.1	1.1
Alcohol (%)					
	Abstainer	24.4	25.7	28.7	29.9
	0-7 drinks/week	62.4	60.8	58.4	55.5

	>7 drinks/week	13.2	13.5	12.9	14.6
Smoking (%)					
	Current	9.6	8.8	7.8	9.3
	Former	38.7	33.7	31.8	32.4
	Never	51.7	57.5	60.4	58.3
Education (%)					
	< High school	5.4	5.8	5.9	5.6
	High school grad and some college	62.2	65.2	65.2	64
	College grad	15.9	15.3	15.6	16.1
	Postgraduate	16.6	13.8	13.4	14.3
Live births (%)					
	None	9.9	8.8	8.5	8.6
	1	7.4	6.6	7	7.2
	2	24.6	23.8	23.1	22.7
	3	25.2	25.4	25.3	25.5
	≥ 4	32.9	35.4	36.1	36
Age at menopause (%)					
	< 40	14.4	14	13.3	13.3
	40-44	14.3	13.6	14.5	13.4
	45-59	23.9	23.9	23.2	23.2
	50-54	36.2	37.3	37.9	38.2
	≥55	11.2	11.3	11.2	12
Family history of breast cancer (%)					
	No	85	85.7	84.6	84.3
	Yes, immediate female	13.9	13.4	14.1	14.5
	Male only	0.2	0.1	0.2	0.1
Bilateral oophorectomy (%)					
	No	90.3	88.9	88.5	88.3
	Yes	9.8	11.1	11.5	11.7

Hours of vigorous PA per week (%)				
None	12.2	14.5	15.6	19.4
< 1	16.3	18.9	19.4	19.7
1	11.3	12.4	11.7	12.3
2	17.7	16.5	17	16.1
3	17.8	16.7	17	14.9
≥4	24.9	21	19.3	17.6
Non-whole/refined grains (oz/day; mean ± SD)	3.51 ± 1.58	3.50 ± 1.52	4.10 ± 1.60	5.66 ± 2.19
Tomatoes (cups/day; mean ± SD)	0.38 ± 0.23	0.36 ± 0.22	0.41 ± 0.24	0.56 ± 0.42
Other vegetables (cups/day; mean ± SD)	1.14 ± 0.64	0.93 ± 0.51	0.90 ± 0.48	0.98 ± 0.52
Cruciferous vegetables (cups/day; mean ± SD)	0.30 ± 0.28	0.27 ± 0.24	0.27 ± 0.24	0.30 ± 0.28
Cheese (cups/day; mean ± SD)	0.23 ± 0.18	0.24 ± 0.18	0.32 ± 0.21	0.60 ± 0.40
Yogurt (cups/day; mean ± SD)	0.25 ± 0.30	0.10 ± 0.15	0.07 ± 0.12	0.06 ± 0.12
Fish/shellfish high in ω-3 fatty acids (oz/day; mean ± SD)	0.19 ± 0.21	0.14 ± 0.15	0.13 ± 0.15	0.17 ± 0.22
Fish/shellfish low in ω-3 fatty acids (oz/day; mean ± SD)	0.65 ± 0.63	0.45 ± 0.37	0.43 ± 0.35	0.49 ± 0.42
Franks and luncheon meats (oz/day; mean ± SD)	0.14 ± 0.17	0.16 ± 0.17	0.22 ± 0.22	0.42 ± 0.45
Nuts and seeds (oz/day; mean ± SD)	0.63 ± 0.97	0.35 ± 0.48	0.33 ± 0.42	0.38 ± 0.45
Coffee (cups/day; mean ± SD)	3.83 ± 4.22	2.37 ± 2.50	1.86 ± 2.18	1.86 ± 2.29

ERDP: estrogen related dietary pattern; ER: estrogen receptor; BMI: body mass index; HRT: hormone replacement therapy; PA: physical activity; PLCO: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; SD: standard deviation

Table 4.4 Hazard ratios (95% CI) for the relationship between the estrogen-related dietary pattern (ERDP) score and postmenopausal breast cancer in PLCO

	ERDP quartiles				Estimate for continuous ERDP score[a], p-trend
	1st	2nd	3rd	4th	
Total breast cancer					
No. of cases	366	392	403	431	
Age-adjusted	1.00 (ref)	1.07 (0.93, 1.24)	1.10 (0.95, 1.26)	1.18 (1.03, 1.36)	1.12 (1.03, 1.20) p=0.005
Age- and-TEI adjusted	1.00 (ref)	1.09 (0.94, 1.25)	1.10 (0.95, 1.27)	1.15 (1.00, 1.33)	1.10 (1.01, 1.18) p=0.02
Multivariable-adjusted[b]	1.00 (ref)	1.08 (0.94, 1.25)	1.10 (0.95, 1.27)	1.14 (0.98, 1.32)	1.09 (1.01, 1.18) p=0.04
Invasive					
No. of cases	280	309	331	348	
Age-adjusted	1.00 (ref)	1.11 (0.94, 1.30)	1.18 (1.00, 1.38)	1.25 (1.07, 1.47)	1.16 (1.07, 1.26) p=0.0006
Age- and-TEI adjusted	1.00 (ref)	1.12 (0.95, 1.32)	1.18 (1.01, 1.38)	1.21 (1.03, 1.43)	1.14 (1.04, 1.24) p=0.003
Multivariable-adjusted[b]	1.00 (ref)	1.12 (0.95, 1.31)	1.18 (1.01, 1.39)	1.20 (1.02, 1.42)	1.13 (1.04, 1.24) p=0.005
ER+					
No. of cases	246	275	274	302	
Age-adjusted	1.00 (ref)	1.12 (0.94, 1.33)	1.11 (0.93, 1.32)	1.23 (1.04, 1.46)	

						1.15 (1.05, 1.26) p=0.003
	Age- and-TEI adjusted	1.00 (ref)	1.13 (0.95, 1.35)	1.11 (0.94, 1.32)	1.20 (1.01, 1.42)	1.13 (1.03, 1.24) p=0.01
	Multivariable-adjusted[b]	1.00 (ref)	1.13 (0.95, 1.34)	1.11 (0.94, 1.32)	1.19 (0.99, 1.41)	1.13 (1.02, 1.24) p=0.02
ER-						
	No. of cases	45	41	55	48	
	Age-adjusted	1.00 (ref)	0.92 (0.60, 1.40)	1.21 (0.82, 1.80)	1.06 (0.71, 1.59)	1.09 (0.87, 1.35) p=0.46
	Age- and-TEI adjusted	1.00 (ref)	0.93 (0.61, 1.43)	1.22 (0.82, 1.81)	1.01 (0.66, 1.53)	1.06 (0.85, 1.32) p=0.63
	Multivariable-adjusted[b]	1.00 (ref)	0.94 (0.61, 1.44)	1.24 (0.83, 1.84)	1.04 (0.68, 1.59)	1.07 (0.85, 1.35) p=0.54
Person-years accumulated		74,615	74,375	74,932	74,468	

ERDP: estrogen related dietary pattern; TEI: total energy intake; ER: estrogen receptor; BMI: body mass index; HRT: hormone replacement therapy; PA: physical activity; PLCO: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

[a]HR corresponds to 1-unit increase in ERDP score.

[b]Includes adjustment for age, TEI, BMI, BMI at age 20, HRT, alcohol use, education, bilateral oophorectomy, parity, age at menopause, PA, race/ethnicity, recruitment center, and family history of breast cancer.

Table 4.5 Hazard ratios (95% CI) for the relationship between the estrogen-related dietary pattern (ERDP) score and postmenopausal breast cancer within strata of estrogen-related risk factors in PLCO[a]

	ERDP quartiles				p interaction[b]
	1st	2nd	3rd	4th	
HRT use at baseline					0.64
No	1.00 (ref)	1.20 (0.96, 1.50)	1.27 (1.02, 1.59)	1.24 (0.99, 1.56)	
Yes	1.00 (ref)	1.02 (0.85, 1.23)	1.01 (0.84, 1.21)	1.07 (0.88, 1.29)	
BMI (kg/m^2)					0.59
18.5-29.9	1.00 (ref)	1.09 (0.92, 1.28)	1.14 (0.97, 1.34)	1.20 (1.02, 1.43)	
≥30	1.00 (ref)	1.12 (0.83, 1.51)	1.01 (0.75, 1.35)	0.99 (0.75, 1.32)	
Alcohol consumption					0.90
<1 drink/week	1.00 (ref)	1.11 (0.92, 1.34)	1.15 (0.96, 1.38)	1.15 (0.95, 1.39)	
≥1 drinks/week	1.00 (ref)	1.05 (0.84, 1.32)	1.03 (0.82, 1.30)	1.14 (0.90, 1.44)	
Parity					0.58
Nulliparous	1.00 (ref)	1.24 (0.77, 2.00)	1.44 (0.90, 2.28)	1.45 (0.91, 2.32)	
Parous	1.00 (ref)	1.06 (0.91, 1.24)	1.07 (0.92, 1.24)	1.10 (0.95, 1.28)	
Vigorous PA					0.61
<2 hours/week	1.00 (ref)	1.18 (0.95, 1.47)	1.11 (0.89, 1.38)	1.12 (0.90, 1.40)	
≥2 hours/week	1.00 (ref)	1.01 (0.83, 1.22)	1.10 (0.91, 1.33)	1.17 (0.96, 1.42)	

ERDP: estrogen related dietary pattern; TEI: total energy intake; ER: estrogen receptor; BMI: body mass index; HRT: hormone replacement therapy; PA: physical activity; PLCO: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

[a]Includes adjustment for age, TEI, BMI, BMI at age 20, HRT, alcohol use, education, bilateral oophorectomy, parity, age at menopause, PA, race/ethnicity, recruitment center, and family history of breast cancer.

[b]P-value for the product term of ERDP quartiles with the potential effect modifier.

CHAPTER 5

AN ESTROGEN-RELATED DIETARY PATTERN AND POSTMENOPAUSL BREAST CANCER RISK IN A COHORT OF WOMEN WITH A FAMILY HISTORY OF BREAST CANCER

Mark A. Guinter,[1,2] Dale P. Sandler,[3] Alexander C. McLain,[1] Anwar T. Merchant,[1] Susan E. Steck[1,2]

[1]Department of Epidemiology and Biostatistics; [2]Cancer Prevention and Control Program, Arnold School of Public Health, University of South Carolina, Columbia, SC; [3]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC.

To be submitted

5.1 Abstract

Introduction: The results of previous studies on diet and postmenopausal breast cancer risk have been inconclusive. There is some evidence that dietary patterns developed to correlate with estrogen have positive associations with breast cancer, however, results are mixed. We applied an estrogen-related dietary pattern (ERDP) that was developed in a subsample of the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) to the Sister Study and examine associations with postmenopausal breast cancer. **Methods:** Participants from the Sister Study without a personal history of cancer and who contributed postmenopausal person-time at risk were included in the present analysis. Intakes of non-whole/refined grains, tomatoes, cruciferous vegetables, other vegetables, cheese, yogurt, fish/shellfish, franks/luncheon meats, nuts and seeds, and coffee were measured via food frequency questionnaires and used to calculate the ERDP. Cox proportional hazards models were used to estimate hazard ratios (HR) and 95% confidence intervals (CI) for the association between the ERDP and postmenopausal breast cancer. **Results:** Over 274,308 person-years of follow-up 1,951 incident cases occurred. ERDP was not associated with total, invasive, ER+, or ER- subtypes of breast cancer either as a continuous or categorical variable. The association did not differ across strata of other estrogen-related risk factors. Results were robust to various sensitivity analyses. **Conclusion:** Our investigation did not support previous studies observing an association between an estrogen-derived dietary pattern and postmenopausal breast cancer risk. Null results may be partially explained by higher levels of other breast cancer risk factors, such as a family history of breast cancer within the study population.

5.2 Introduction

Breast cancer accounts for nearly one-third of incident cancer cases among U.S. women and imposes a significant disease burden.[277] Primary prevention may help ease this burden, yet the identification of modifiable lifestyle factors for prevention remains a large gap in translational breast cancer research.[43] Diet represents a commonly studied lifestyle behavior in cancer prevention, however, results from studies of individual dietary components has yielded inconsistent results.[17–19,24] Among dietary factors, only alcohol is recognized by the World Cancer Research Fund and American Institute for Cancer Research to have a probable influence on postmenopausal breast cancer risk, as determined by the observational evidence and biologic plausibility from experimental studies.[24]

It is possible that focusing on a known biologic mechanism while considering the totality of diet and not only individual foods or nutrients may result in the identification of stronger dietary associations with breast cancer. In postmenopausal breast cancer, circulating or urinary estrogen metabolites have been associated with disease risk.[11] Postmenopausal women have low endogenous levels of estrogen, therefore relatively small changes in estrogen resulting from dietary exposures may play a role in breast cancer risk.

Accordingly, we and others have developed dietary patterns based on associations between specific food groups and measured levels of estrogens or estrogen metabolites.[27,32] Results of studies linking these dietary patterns to postmenopausal breast cancer risk have been mixed. In the Nurses' Health Study (NHS), a dietary pattern

based on estradiol (E2) and estrone sulfate was not associated with postmenopausal breast cancer risk.[32] However, the same pattern was applied in a group of Swedish women and a positive association with breast cancer was observed.[27] A second estrogen-related dietary pattern (ERDP) was developed using data from a nested study of participants from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO).[35,278] The ERDP was developed using unconjugated E2 and the ratio of 2- to 16-hydroxylated (2/16 ratio) estrogen metabolites (EM). Fifteen downstream metabolites, including the 2- and 16-pathways, could be quantified using a liquid chromatography-tandem mass spectrometry assay (LC/MS-MS) which has enough sensitivity to detect the low levels present in postmenopausal women.[262] The ERDP was positively associated with postmenopausal breast cancer risk in a time-to-event analysis in the PLCO cohort.[278]

We applied the ERDP in a study population different from the one in which it was developed and examined its association with postmenopausal breast cancer risk. Outcomes included total and invasive breast cancer, as well as estrogen receptor (ER) subtypes of breast cancer. The potential for a differential effect in strata of other estrogen-related risk factors was assessed. We hypothesized the ERDP to be positively associated with breast cancer. We expected to see the strongest associations in ER+ subtypes and among strata of other risk factors where the relative estrogen exposure was lowest.

5.3 Methods

5.3.1 Study Population

The Sister Study, an initiative of the National Institute of Environmental Health Sciences, is a large prospective cohort study to investigate environmental and genetic determinants of breast cancer.[34] A total of 50,884 women aged 35 to 74 who had a sister who was diagnosed with breast cancer were recruited between 2003 and 2009 from all 50 U.S. states and Puerto Rico. Women completed self-administered questionnaires and a computer-assisted telephone interview to ascertain information on demographics and potential risk factors. Breast cancer incidence was assessed annually via a brief study update and a comprehensive follow-up questionnaire was administered every two to three years. The Institutional Review Board of the National Institute of Environmental Health Sciences and the Copernicus Group Institutional Review Board approved the study. In the present analysis, participants were excluded if they had a personal history of cancer (n=2,757), or if they did not contribute any person-time at risk for postmenopausal breast cancer (n=8,004). Participants were further excluded if they reported an extreme caloric intake (<500 or >5,000 kcal/day; n=1,163), had an extreme body mass index (BMI; <15 or >50 kg/m^2; n=68), or if they had missing covariate data (n=1,140); bringing the total analytic sample to 37,752.

5.3.2 Dietary Assessment

Intakes over the prior 12 months, as measured by the 110-item 1998 Block food frequency questionnaire (FFQ) at baseline, were used to calculate the ERDP. Briefly, the ERDP was previously developed using reduced rank regression modeling to identify food

groups that were associated with serum levels of unconjugated E2 and the 2/16 ratio in a nested case-control study of 653 postmenopausal women from PLCO. The resulting ERDP was comprised of non-whole/refined grains, tomatoes, cruciferous vegetables, cheese, fish/shellfish high in ω-3 fatty acids, franks/luncheon meats, nuts and seeds, other vegetables, fish/shellfish low in ω-3 fatty acids, yogurt, and coffee. Intakes of the food groups were centered and scaled, then multiplied by their corresponding model weights which were derived using the PLCO data (**Table 4.1**). The total ERDP score was calculated by summing over the weighted intakes. In the present study, intakes from the FFQ were categorized into food groups using the U.S. Department of Agriculture's Food Patterns Equivalents Database.[279] Additional food groups were created for cruciferous vegetables and coffee. Higher ERDP scores are hypothesized to be positively associated with unconjugated E2 and inversely associated with the 2/16 ratio.

5.3.3 Breast Cancer Ascertainment

Participants were followed until breast cancer diagnosis, death, or end of follow-up. Incident breast cancer cases were ascertained via completion of annual health updates and biennial surveys. Response rates for the surveys were over 94%.[253] Access to medical records was requested after a breast cancer diagnosis was self-reported. Medical record abstraction was used to confirm over 80% of cases and to identify information on treatment and diagnosis, such as ER subtype.[200] The positive predictive value of self-reported breast cancer, invasive cancer, and ER breast cancer was over 90% and therefore self-reported information is used when medical records could not be obtained.[254]

5.3.4 Statistical Approach

Baseline comparisons of participant characteristics by ERDP quartiles were performed using t-tests and chi-square tests for continuous and categorical variables, respectively. Cox proportional hazards models were applied to analyze the relationship between ERDP scores and incident breast cancer events, using age as the time scale variable. The proportional hazards assumption was evaluated using Martingale-based residuals and was not violated by exposure variables or covariates. ERDP scores were categorized into quartiles, with the first quartile set as the referent. The first quartile hypothetically represents diets with an estrogen profile associated with the lowest breast cancer risk (low levels of unconjugated E2 and high 2/16 ratio). The hazard ratios and 95%CIs also were calculated for the continuous ERDP score variable, and the p-value reported as a test for trend. Demographic factors of age (years), education (less than high school; high school and some college; college degree; graduate degree), race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian, other), and total caloric intake (kcal/day) were included in the multivariable-adjusted models for their putative roles as confounders, as identified by a DAG. The remaining covariates included in multivariable-adjusted models were chosen using stepwise model selection with entry/exit criteria of p=0.2, and include hormone replacement therapy (HRT) use (never, former – estrogen + progesterone, former – estrogen only, current – estrogen + progesterone, current – estrogen only, ever – unknown type), baseline body mass index (BMI; kg/m^2), BMI at age 30 kg/m^2), physical activity (metabolic equivalent of task (MET)-hours/week), alcohol consumption (abstainer, ≤1 drink/day, >1 drink/day), number of first degree relatives with history of breast cancer, age at menarche, age at

menopause, parity, and prior hysterectomy (yes, no). Oral contraceptive (OC) use, age at first birth, bilateral oophorectomy, and smoking status were considered but were not retained after stepwise selection in order to improve model efficiency and reduce the potential for over adjustment. Models were used with total breast cancer, invasive only, and ER subtypes among invasive cases as outcomes. A competing risk model assessed a differential association for the ERDP on ER+ and ER- subtypes using a Wald test for heterogeneity in the stratified Lunn-McNeil approach.[259] Effect modification by baseline HRT use (yes, no), BMI (18.5-29.9 kg/m^2, ≥30 kg/m^2), alcohol consumption (abstainer, ≤1 drink/day, >1 drink/day), parity (nulliparous, parous), and whether or not participants met the Physical Activity Guidelines for Americans (PAG; <500 MET-min/week, ≥500 MET-min/week).[280] Soy products are not a component of the ERDP. However, soy foods (e.g., tofu, tempeh, soy milk, and other soy substitutes) may modify dietary influences on breast cancer due to their high phytoestrogen content.[167] Additional questions on soy food intake were added to the Sister Study FFQ. Therefore, the association between the ERDP and breast cancer was assessed in strata of soy food consumption (non-consumer, >0 to 4.9 g/day, ≥5 g/day).

Multiple sensitivity analyses were conducted. The first set of sensitivity analyses assessed changing parameters in the model. If BMI lies on the causal pathway between the ERDP and postmenopausal breast cancer, there is the potential for mediator bias. Therefore, a sensitivity analysis was conducted by removing BMI as a covariate. Because hormone receptor status was not routinely obtained for *in situ* cases during the study period, ER receptor subtype analyses were limited to invasive cases. However, a sensitivity analysis was carried out including *in situ* cases. A second set of sensitivity

analyses examined the relationship between the ERDP and breast cancer in population subgroups. Minority population recruitment was prioritized at later stages in the recruitment process and therefore minority women have slightly shorter average follow-up time. Thus, the relationship was investigated when restricting to non-Hispanic Whites. In another subgroup analysis, participants who contributed ≤12 months of follow-up were excluded to minimize the possibility of reverse causality. The final subgroup analysis was restricted to women with no more than one full sister having a history of breast cancer, as effects of lifestyle factors such as diet may be harder to detect in women with a strong inherited risk. All statistical tests were two-sided at $\alpha=0.05$, with the exception of interaction p-values which were considered statistically significant at $p<0.10$. All analyses were conducted using SAS 9.4 (SAS Institute, Cary, NC).

5.4 Results

There were 1,951 incident cases of postmenopausal breast cancer over 274,308 person-years of follow-up. Among the 1,484 invasive cases, 1,098 and 199 were ER+ and ER-, respectively. The mean ± standard deviation (SD) ERDP score was -0.05 ± 0.71 with a range of -8.32 to 4.67. Average ERDP scores among women who were diagnosed with incident breast cancer during follow-up were not significantly different than women who were not diagnosed (-0.06 vs 0.04, respectively; p=0.28). Baseline characteristics across strata of ERDP quartiles are shown in **Table 5.1**. Women in the fourth ERDP quartile were typically younger, had a higher BMI at baseline and at age 30, consumed more calories but less alcohol, and were more likely to have never used HRT or be former users of an estrogen + progesterone formula of HRT. These women also were

more likely to be non-Hispanic White or Hispanic. Intakes of non-whole/refined grains, tomatoes, cheese, franks and luncheon meats were highest among participants in the fourth ERDP quartile. Conversely, participants in the first ERDP quartile consumed higher amounts of other vegetables, yogurt, fish/shellfish, nuts and seeds, and coffee.

In multivariable models comparing the highest ERDP quartile with the first (**Table 5.2**), or diets with the most estrogenic potential compared to the least potential, no association was observed for total (HR: 0.98; 95%CI: 0.86, 1.11; p_{trend}=0.70), invasive (HR: 0.96; 95%CI: 0.83, 1.11; p_{trend}=0.41), invasive ER+ (HR: 0.91; 95%CI: 0.77, 1.07; p_{trend}=0.12), or invasive ER- (HR: 1.23; 95%CI: 0.82, 1.84; p_{trend}=0.17) breast cancer. Results from a competing risk model indicated there was no differential effect of the ERDP on ER+ and ER- subtypes (p=0.18; data not shown). **Table 5.3** shows evidence for potential effect modification between the ERDP and total breast cancer by alcohol consumption (p=0.03), parity (p=0.03), and whether or not participants met the criteria of 500 MET-min/week from the PAG (p=0.07). In the fourth ERDP quartile estimates of association were highest among participants who consumed ≥1 alcohol drink per day (HR: 1.18; 95%CI: 0.83, 1.69), were nulliparous (HR: 1.34; 95%CI: 1.00, 1.80), or who exercised for 500 MET-min/week (HR: 1.13; 95%CI: 0.95, 1.34), although CIs included the null value. There was no evidence of effect modification by HRT use at baseline (p=0.34) or BMI status at baseline (p=0.44).

No association for the relationship between the ERDP and postmenopausal breast cancer was observed in sensitivity models with different model parameters (**Table 5.4**) or within different population subgroups (**Table 5.5**). Similarly, there was no association

between the ERDP and total breast cancer within strata of soy food consumption levels (**Table 5.6**).

5.5 Discussion

A dietary pattern derived in PLCO to correlate with a high-risk estrogen profile (high unconjugated E2, low 2/16 ratio) was applied in the Sister Study, prospective cohort of women with a family history of breast cancer. Results from the time-to-event analysis showed no association for the ERDP with total, invasive, or ER subtypes of postmenopausal breast cancer. Participants in the fourth ERDP quartile, who were suspected to have the most estrogenic potential from their diets, did not experience greater risk of postmenopausal breast cancer compared to individuals in the first quartile.

Prior application of the ERDP in PLCO participants yielded a 9% increase in risk of total postmenopausal breast cancer for a 1-unit increase in ERDP score (**Table 4.4**). Furthermore, a 20% and 19% increase in risk was observed for invasive (HR: 1.20; 95%CI: 1.01, 1.42) and ER+ cases (HR: 1.19; 95%CI: 0.99, 1.41) when comparing participants in the fourth quartile of ERDP scores to the first. In addition to the potential for a true null association, differences in results from the PLCO and present analyses may be due to differences in the FFQs used to measure dietary intakes as well as due to characteristics of the populations with respect to other breast cancer risk factors. As the FFQs used in the present analysis and the PLCO cohort are close-ended, the intakes of the food groups used in the ERDP are dependent upon quantity and description of the line items containing those foods. It is likely that these differences resulted in different distributions of the scores between the two populations, as evident by the range of -8.23

to 4.67 in the Sister Study compared to a range of -4.52 to 6.58 in PLCO participants. Furthermore, there was a stepwise decrease in person-time contributed across the ERDP quartiles. Participants in the fourth quartile, which was hypothesized to be most strongly associated with breast cancer, were younger and contributed the least amount of person-time (**Table 5.2**). It is possible that some of these participants with the most estrogenic diets did not contribute enough postmenopausal time at risk in order for a dietary influence on estrogen to take effect and result in incident breast cancer, although the relative difference time contributed per person was minimal.

It is also possible that the relative prevalence of other breast cancer risk factors between the two populations contributed to different results. A prominent difference between the two study populations is the presence of a family history of breast cancer for all Sister Study participants compared to a much lower proportion in PLCO (discussed further in section 7.1). Women with at least one first-degree relative have roughly two times the risk of developing postmenopausal breast cancer.[281] Therefore, it is plausible that the increase in risk associated with a family history of breast cancer may render a dietary association more difficult to detect, although evidence for a modifying effect by family history in dietary studies is limited.[227] In addition to inherited risk, Sister Study participants had higher levels of other risk factors compared to PLCO. For example, participants in the present analysis were less educated, more likely to drink alcohol, had higher prevalence of past hysterectomy, and were more likely to experience the onset of menopause after the age of 55. Harris et al. also cited lower prevalence of other risk factors in the Swedish women as a potential explanation as to why they observed an association with the estrogen correlated dietary pattern which was not seen in the NHS.[27]

Although significant p-values were observed for modifying effects of alcohol, parity, and PA on the ERDP's association with breast cancer, all of the estimates of association were null and there were no clear trends in the quartiles. There was a suggestion of higher risk for the ERDP among nulliparous women (HR_{Q4vsQ1}: 1.34; 95%CI: 1.00, 1.80). It is possible that some women did not have children as a result of low or imbalanced hormones affecting their fertility,[264] which would potentially explain why a hormone-related dietary influence was detected and the association observed in the fourth ERDP quartile for nulliparous women.

An important take away from these results, along with those from Harris et al., are the implications of adapting dietary patterns, particularly *a posteriori* patterns, across study populations with different dietary assessment tools. Intakes of commonly consumed foods in one population may be measured differently or be completely absent in another population, thereby having an influence on the scoring of the dietary pattern. For example, in the NHS estrogen correlated dietary pattern, pizza was a contributing food group. However, pizza is not commonly consumed in Sweden, so Harris et al. dichotomized pizza consumption into consumers versus not, rather than use the servings/day intake as in the NHS. It is likely that an estrogen-related dietary pattern derived from RRR in a population with different diets, such as in Asia, could result in a different pattern of food groups with little to no overlap. An open-ended dietary questionnaire would allow for full ascertainment of foods that may influence estrogen metabolism, which was not applicable in the present population or the one in which the ERDP was derived.

There are some other minor limitations in addition to the challenges associated with applying dietary patterns across different study populations. As with all self-reported data, there is the potential for inaccurate reporting, which could result in misclassification of dietary intakes or other risk factors. There also is the potential for selection bias due to loss to follow up, although response rates in this highly motivated cohort were high (>90%). Our study was primarily comprised of non-Hispanic White women, therefore generalizability may be limited, although breast cancer incidence is highest in this population.[277] Low numbers of ER- cases may have made it difficult to detect a potential association and assess a differential effect of the ERDP on ER subtypes. Strengths in the present analysis include a prospective study population with complete information on known confounders for the relationship between diet and breast cancer. The large sample size allowed for relatively small effects to be detected, therefore it is unlikely that a lack of power contributed to the null results.

In conclusion, the ERDP, which was based on an estrogen profile hypothesized to increase breast cancer risk, was not associated with risk of postmenopausal breast cancer. All participants in the present study population had a family history of breast cancer, therefore the inherited risk and high prevalence of other breast cancer risk factors may have contributed to the lack of an association. Our analysis highlights the difficulties in comparing *a posteriori* dietary patterns across populations, and suggests the importance of considering dietary measurement tools when interpreting results from dietary investigations.

5.6 Tables

Table 5.1 Sister Study population characteristics across quartiles of the estrogen-related dietary pattern (ERDP) score

		ERDP Quartile (score range)			
		1st (-8.230 -0.444)	2nd (-0.445, -0.058)	3rd (-0.059, 0.347)	4th (0.348, 4.670)
n		9,438	9,438	9,438	9,438
Breast cancer cases	Total	515	472	481	483
	Invasive	393	372	349	370
	Invasive ER+	308	272	247	271
	Invasive ER-	46	45	53	55
ERDP score (mean ± SD)		-0.90 ± 0.48	-0.24 ± 0.11	0.13 ± 0.12	0.82 ± 0.47
Age (mean ± SD)		58.3 ± 7.3	58.0 ± 7.5	57.1 ± 7.8	56.0 ± 7.6
BMI (kg/m^2; mean ± SD)		27.0 ± 5.5	27.3 ± 5.6	28.0 ± 6.1	28.7 ± 6.6
BMI at age 30 (kg/m^2; mean ± SD)		22.7 ± 3.3	22.7 ± 3.3	23.0 ± 3.5	23.4 ± 4.0
Total caloric intake (kcal/day)		1,600 ± 590	1,440 ± 533	1,527 ± 535	1,910 ± 650
MET-hours/week (mean ± SD)		53.3 ± 31.9	51.1 ± 31.0	49.6 ± 30.9	49.7 ± 31.0
Age at menarche (mean ± SD)		12.6 ± 1.5	12.6 ± 1.5	12.6 ± 1.5	12.7 ± 1.6
Age at menopause (mean ± SD)		50.3 ± 5.8	50.1 ± 5.9	49.9 ± 6.1	49.8 ± 5.9
Number of relatives with family history (mean ± SD)		1.27 ± 0.59	1.28 ± 0.59	1.27 ± 0.58	1.26 ± 0.56
Parity (mean ± SD)		1.95 ± 1.35	2.01 ± 1.36	2.01 ± 1.37	1.96 ± 1.36
Nulliparous (%)		18.2	16.6	17.2	18.5
HRT status (%)					
	Never	44.5	45.7	47.9	51.8
	Former - Estrogen+Progesterone	24.2	22.7	21.7	19.3

	Former - Estrogen only	16.8	17.0	15.5	14.6
	Former - unknown what type	2.6	2.6	2.8	2.9
	Current - Estrogen+Progesterone	5.2	4.9	5.1	4.6
	Current - Estrogen only	6.7	7.1	7.0	6.8
Race/ethnicity (%)					
	White, Non-Hispanic	85.7	84.8	85.1	86.5
	Black, Non-Hispanic	8.2	8.0	7.9	5.5
	Hispanic	2.8	4.3	4.5	5.1
	Asian	0.9	0.7	0.4	0.3
	Other	2.4	2.2	2.1	2.6
Alcohol (%)					
	Abstainer	17.7	18.5	19.8	21.1
	≤1 drink/day	70.9	70.0	69.7	67.1
	>1 drink/day	11.4	11.5	10.5	11.8
Smoking (%)					
	Current	7.9	7.8	7.6	8.0
	Former	41.7	38.5	34.9	35.1
	Never	50.4	53.7	57.6	56.9
Education (%)					
	< HS	0.9	1.1	1.2	1.2
	HS grad and some college	32.4	35.0	36.2	34.8
	College grad	39.5	39.1	39.6	41.3
	Postgraduate	27.2	24.8	23.0	22.7
Hysterectomy					
	No	65.6	65.6	65.1	66.9
	Yes	34.4	34.4	34.9	33.1
Non-whole/refined grains (oz/day; mean ± SD)		2.32 ± 1.26	2.39 ± 1.24	2.82 ± 1.34	3.94 ± 1.98

Tomatoes (cups/day; mean ± SD)	0.24 ± 0.17	0.23 ± 0.17	0.25 ± 0.19	0.35 ± 0.27
Other vegetables (cups/day; mean ± SD)	0.63 ± 0.51	0.44 ± 0.33	0.40 ± 0.29	0.45 ± 0.31
Cruciferous vegetables (cups/day; mean ± SD)	0.23 ± 0.27	0.20 ± 0.25	0.22 ± 0.25	0.29 ± 0.40
Cheese (cups/day; mean ± SD)	0.26 ± 0.21	0.28 ± 0.21	0.36 ± 0.24	0.66 ± 0.41
Yogurt (cups/day; mean ± SD)	0.23 ± 0.28	0.10 ± 0.14	0.08 ± 0.12	0.06 ± 0.11
Fish/shellfish high in ω-3 fatty acids (oz/day; mean ± SD)	0.18 ± 0.23	0.14 ± 0.17	0.13 ± 0.17	0.15 ± 0.22
Fish/shellfish low in ω-3 fatty acids (oz/day; mean ± SD)	0.58 ± 0.63	0.40 ± 0.37	0.38 ± 0.34	0.42 ± 0.41
Franks and luncheon meats (oz/day; mean ± SD)	0.38 ± 0.32	0.41 ± 0.33	0.51 ± 0.37	0.78 ± 0.59
Nuts and seeds (oz/day; mean ± SD)	2.11 ± 2.26	1.31 ± 1.32	1.14 ± 1.17	1.20 ± 1.20
Coffee (cups/day; mean ± SD)	2.19 ± 1.71	1.58 ± 1.43	1.19 ± 1.28	1.02 ± 1.25

BMI: body mass index; ER: estrogen receptor; ERDP: estrogen related dietary pattern; HRT: hormone replacement therapy; HS: high school; MET: metabolic equivalent of task.

Table 5.2 Hazard ratios (95% CI) for the relationship between the estrogen related dietary pattern (ERDP) score and postmenopausal breast cancer in the Sister Study

	ERDP quartiles				
	1st	2nd	3rd	4th	Estimate for continuous ERDP score[a]
Total breast cancer					
No. of cases	515	472	481	483	
Age-adjusted	1.00 (ref)	0.93 (0.82, 1.05)	0.98 (0.87, 1.11)	1.03 (0.91, 1.17)	1.02 (0.95, 1.08) p=0.60
Age- and TEI-adjusted	1.00 (ref)	0.94 (0.83, 1.06)	0.99 (0.87, 1.12)	1.00 (0.88, 1.14)	1.00 (0.94, 1.07) p=0.93
Multivariable-adjusted[b]	1.00 (ref)	0.93 (0.82, 1.05)	0.97 (0.85, 1.10)	0.98 (0.86, 1.11)	0.99 (0.93, 1.06) p=0.70
Invasive					
No. of cases	393	372	349	370	
Age-adjusted	1.00 (ref)	0.96 (0.83, 1.10)	0.94 (0.81, 1.09)	1.04 (0.90, 1.20)	1.02 (0.95, 1.09) p=0.65
Age- and TEI-adjusted	1.00 (ref)	0.98 (0.85, 1.13)	0.95 (0.82, 1.10)	1.00 (0.87, 1.16)	1.00 (0.93, 1.07) p=0.92
Multivariable-adjusted[b]	1.00 (ref)	0.96 (0.83, 1.10)	0.92 (0.79, 1.06)	0.96 (0.83, 1.11)	0.97 (0.90, 1.04) p=0.41
Invasive ER+					
No. of cases	308	272	247	271	
Age-adjusted	1.00 (ref)	0.89 (0.76, 1.05)	0.85 (0.72, 1.01)	0.99 (0.84, 1.16)	0.98 (0.90, 1.07) p=0.71

	Age- and TEI-adjusted	1.00 (ref)	0.91 (0.78, 1.08)	0.86 (0.73, 1.02)	0.95 (0.80, 1.12)	0.96 (0.89, 1.05) p=0.37
	Multivariable-adjusted[b]	1.00 (ref)	0.90 (0.76, 1.06)	0.84 (0.71, 0.99)	0.91 (0.77, 1.07)	0.94 (0.86, 1.02) p=0.12
Invasive ER-						
	No. of cases	46	45	53	55	
	Age-adjusted	1.00 (ref)	0.99 (0.66, 1.49)	1.19 (0.80, 1.77)	1.26 (0.85, 1.87)	1.17 (0.96, 1.43) p=0.11
	Age- and TEI-adjusted	1.00 (ref)	1.00 (0.66, 1.51)	1.20 (0.81, 1.78)	1.24 (0.83, 1.85)	1.16 (0.95, 1.42) p=0.15
	Multivariable-adjusted[b]	1.00 (ref)	0.99 (0.66, 1.50)	1.19 (0.80, 1.76)	1.23 (0.82, 1.84)	1.16 (0.94, 1.42) p=0.17
Person-years accumulated		69,826	69,274	68,102	67,106	

ER: estrogen receptor; ERDP: estrogen related dietary pattern; TEI: total energy intake.
aHR corresponds to 1-unit increase in ERDP score.
bIncludes adjustment for age, TEI, BMI, BMI at age 30, HRT, race/ethnicity, alcohol use, number of family members with a history of breast cancer, age at menarche, age at menopause, parity, and hysterectomy.

Table 5.3 Hazard ratios (95% CI) for the relationship between the estrogen related dietary pattern (ERDP) score and postmenopausal breast cancer within strata of estrogen-related risk factors in the Sister Study[a]

	ERDP quartiles				
	1st	2nd	3rd	4th	p interaction[b]
HRT use at baseline					0.34
No	1.00 (ref)	0.89 (0.77, 1.02)	0.96 (0.84, 1.10)	0.94 (0.81, 1.08)	
Yes	1.00 (ref)	1.14 (0.84, 1.55)	1.03 (0.75, 1.41)	1.28 (0.94, 1.74)	
BMI (kg/m^2)					0.44
<30	1.00 (ref)	0.92 (0.79, 1.07)	1.02 (0.88, 1.18)	0.96 (0.82, 1.12)	
≥30	1.00 (ref)	0.92 (0.73, 1.16)	0.85 (0.68, 1.07)	1.00 (0.80, 1.25)	
Alcohol consumption					0.03
Abstainer	1.00 (ref)	1.14 (0.85, 1.54)	1.07 (0.80, 1.45)	0.90 (0.66, 1.22)	
<1 drink/week	1.00 (ref)	0.81 (0.70, 0.95)	0.88 (0.76, 1.02)	0.98 (0.84, 1.14)	
≥1 drinks/week	1.00 (ref)	1.29 (0.93, 1.80)	1.37 (0.98, 1.92)	1.18 (0.83, 1.69)	
Parity					0.03
Nulliparous	1.00 (ref)	1.03 (0.76, 1.41)	0.93 (0.68, 1.28)	1.34 (1.00, 1.80)	
Parous	1.00 (ref)	0.90 (0.79, 1.03)	0.97 (0.84, 1.11)	0.91 (0.79, 1.04)	
Meets Physical Activity Guidelines					0.07
<500 MET-min/week	1.00 (ref)	0.94 (0.78, 1.13)	0.91 (0.76, 1.10)	0.84 (0.70, 1.02)	
≥500 MET-min/week	1.00 (ref)	0.90 (0.76, 1.07)	1.01 (0.85, 1.20)	1.13 (0.95, 1.34)	

BMI: body mass index; ERDP: estrogen related dietary pattern; HRT: hormone replacement therapy.
aIncludes adjustment for age, TEI, BMI, BMI at age 30, HRT, race/ethnicity, alcohol use, number of family members with a history of breast cancer, age at menarche, age at menopause, parity, and hysterectomy.
bP-value for the product term of ERDP quartiles with the potential effect modifier.

Table 5.4 Sensitivity analyses with different model parameters

	ERDP quartiles				Estimate for continuous ERDP score[a]
	1st	2nd	3rd	4th	
ER+ including non-invasive cases					
No. of cases	383	340	335	337	
	1.00 (ref)	0.90 (0.78, 1.04)	0.91 (0.79, 1.06)	0.92 (0.79, 1.07)	0.95 (0.88, 1.03) p=0.19
ER+ including non-invasive cases					
No. of cases	63	59	71	71	
	1.00 (ref)	0.95 (0.67, 1.36)	1.18 (0.84, 1.67)	1.21 (0.86, 1.72)	1.11 (0.93, 1.33) p=0.25
Assess BMI mediator bias[c]					
No. of cases	515	472	481	483	
	1.00 (ref)	0.94 (0.83, 1.06)	0.99 (0.88, 1.12)	1.02 (0.89, 1.15)	1.01 (0.95, 1.08) p=0.81

BMI: body mass index; ER: estrogen receptor; ERDP: estrogen related dietary pattern.
[a]HR corresponds to 1-unit increase in ERDP score.
[b]Includes adjustment for age, TEI, BMI, BMI at age 30, HRT, race/ethnicity, alcohol use, number of family members with a history of breast cancer, age at menarche, age at menopause, parity, and hysterectomy.
[c]Includes adjustment for all variables in "b" except for BMI.

Table 5.5 Sensitivity analyses among different analytic population subgroups[a]

	n	ERDP quartiles 1st	2nd	3rd	4th	Estimate for continuous ERDP score[b]
Restricting to non-Hispanic Whites	32,282					
No. of cases		453	416	429	416	
		1.00 (ref)	0.92 (0.81, 1.06)	0.97 (0.85, 1.11)	0.96 (0.84, 1.10)	0.99 (0.93, 1.05) p=0.70
Excluding participants with ≤12 months follow-up	37,369					
No. of cases		489	439	455	452	
		1.00 (ref)	0.91 (0.80, 1.03)	0.97 (0.85, 1.10)	0.97 (0.85, 1.11)	0.98 (0.91, 1.04) p=0.48
Excluding participants with >1 full family member with breast cancer	27,828					
No. of cases		337	300	308	318	
		1.00 (ref)	0.90 (0.77, 1.06)	0.95 (0.81, 1.11)	1.00 (0.85, 1.17)	0.99 (0.92, 1.08) p=0.89

ERDP: estrogen related dietary pattern.

[a]Includes adjustment for age, TEI, BMI, BMI at age 30, HRT, race/ethnicity, alcohol use, number of family members with a history of breast cancer, age at menarche, age at menopause, parity, and hysterectomy.

[b]HR corresponds to 1-unit increase in ERDP score.

Table 5.6 Relationship between the estrogen related dietary pattern (ERDP) and postmenopausal breast cancer across strata of soy food consumption[a]

		Consumption of soy foods		
	ERDP Quartile	Non-consumers (0 g/day)	Low consumers (>0 to 4.9 g/day)	High consumers (≥5 g/day)
Total (no. of cases)		351	992	608
	1st	1.00 (ref)	1.00 (ref)	1.00 (ref)
	2nd	1.02 (0.76, 1.38)	0.95 (0.79, 1.13)	0.84 (0.67, 1.05)
	3rd	0.81 (0.60, 1.11)	0.98 (0.82, 1.17)	1.06 (0.85, 1.32)
	4th	0.92 (0.67, 1.25)	1.00 (0.83, 1.20)	1.01 (0.80, 1.26)
Invasive (no. of cases)		266	753	465
	1st	1.00 (ref)	1.00 (ref)	1.00 (ref)
	2nd	0.97 (0.69, 1.37)	0.99 (0.81, 1.21)	0.90 (0.70, 1.15)
	3rd	0.83 (0.59, 1.19)	0.93 (0.76, 1.14)	0.97 (0.76, 1.25)
	4th	0.88 (0.61, 1.25)	0.99 (0.80, 1.22)	0.99 (0.76, 1.28)
Invasive ER+ (no. of cases)		185	572	341
	1st	1.00 (ref)	1.00 (ref)	1.00 (ref)
	2nd	0.88 (0.60, 1.35)	0.96 (0.76, 1.22)	0.81 (0.61, 1.08)
	3rd	0.78 (0.51, 1.19)	0.92 (0.72, 1.16)	0.77 (0.57, 1.04)
	4th	0.77 (0.50, 1.17)	0.98 (0.77, 1.25)	0.90 (0.67, 1.21)
Invasive ER- (no. of cases)		35	98	66
	1st	1.00 (ref)	1.00 (ref)	1.00 (ref)
	2nd	0.91 (0.35, 2.38)	0.98 (0.54, 1.78)	1.08 (0.53, 2.21)
	3rd	0.77 (0.29, 2.09)	1.16 (0.65, 2.05)	1.55 (0.79, 3.03)
	4th	1.21 (0.46, 3.19)	1.18 (0.66, 2.12)	1.36 (0.68, 2.72)

ER: estrogen receptor; ERDP: estrogen related dietary pattern.
aIncludes adjustment for age, TEI, BMI, BMI at age 30, HRT, race/ethnicity, alcohol use, number of family members with a history of breast cancer, age at menarche, age at menopause, parity, and hysterectomy.

CHAPTER 6

AN ESTROGEN-RELATED LIFESTYLE SCORE IS ASSOCIATED WITH RISK OF POSTMENOPAUSAL BREAST CANCER IN THE PLCO COHORT.

Mark A. Guinter,[1,2] Alexander C. McLain,[1] Anwar T. Merchant,[1] Dale P. Sandler,[3] Susan E. Steck[1,2]

[1]Department of Epidemiology and Biostatistics; [2]Cancer Prevention and Control Program, Arnold School of Public Health, University of South Carolina, Columbia, SC; [3]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC.

To be submitted for publication in *Breast Cancer Research and Treatment*

6.1 Abstract

Lifestyle factors have been associated with estrogen metabolism, which has a strong mechanistic role in the development of postmenopausal breast cancer. We aimed to investigate the combined effect of estrogen-related lifestyle factors on postmenopausal breast cancer risk using data from 27,153 women enrolled in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. We created an estrogen-related lifestyle score (ERLS) by incorporating a previously developed measure of estrogenic diet, alcohol intake, body mass index (BMI), and physical activity. The scores ranged from 0-6 with alcohol and BMI accounting for higher weights than the other factors. To evaluate the preventive possibilities of a low estrogen-related lifestyle, and to keep the direction of the score consistent with other published lifestyle scores, higher scores were set to correspond with potentially lower estrogenic lifestyle. The association between the ERLS and incident breast cancer was examined using Cox proportional hazards models. Participants with an ERLS of 4 or ≥5 had a 23% (HR: 0.77; 95%CI: 0.67-0.89) and 34% (HR: 0.66; 95%CI: 0.56-0.78) lower risk of breast cancer, respectively, compared to those with an ERLS ≤2 after multivariable adjustment. Estimates were similar when restricting to invasive cases or estrogen receptor positive subtypes. No single lifestyle component appeared to drive the association. Our findings suggest that the combined effect of a lifestyle characterized by a low estrogenic diet, low alcohol consumption, low body weight, and high levels of physical activity are associated with a reduction in postmenopausal breast cancer risk, possibly through an influence on estrogen metabolism.

6.2 Introduction

Breast cancer is the second most diagnosed cancer worldwide.[282] In the US, an estimated 250,000 new cases of breast cancer will be diagnosed in 2017, accounting for approximately 30% of all cancer diagnoses in women.[277] Over two-thirds of breast cancers occur in post-menopausal women over the age of 55.[2] Although many well-established risk factors for postmenopausal breast cancer have been identified, not all represent opportunities for primary prevention to help lessen this burden.

There is sufficient evidence to link several lifestyle factors to the development of postmenopausal breast cancer.[2,13] Both sides of the energy balance equation - excess intake in the form of adiposity and greater energy expenditure in the form of physical activity (PA) - show evidence of a positive and negative association with breast cancer, respectively.[2,24] Consumption of alcohol increases breast cancer risk.[2,24] Although evidence of a dietary association with breast cancer is less robust, it is still suggestive.[18,44] Lifestyle factors often cluster together in individuals who adopt healthy or unhealthy lifestyles, so it can be advantageous to study lifestyle factors using a combined lifestyle score.[245] A handful of studies have used indices to assess modifiable lifestyle factors as one aggregate score and have reported consistent, yet moderate, inverse associations between a healthy lifestyle and breast cancer.[14–16,246–250] Previous lifestyle indices were based on adherence to cancer prevention guidelines,[14,250] included behaviors specific to a study population,[15] or were simply based on what is thought to constitute healthy behaviors.[16,246]

Consideration of a disease mechanism in the development of a lifestyle score may help to identify stronger associations than previous studies. In the case of postmenopausal

breast cancer, the prominent influence of estrogen exposure on mammary carcinogenesis is well-documented.[4] Regarding modifiable lifestyle behaviors, increased adiposity and consumption of alcohol are positively associated with estrogen,[140] whereas PA is inversely associated with estrogen;[212] all of which are associated with breast cancer risk.[2] There is recent evidence of dietary patterns developed to correlate with estrogen levels that were subsequently associated with postmenopausal breast cancer risk in some studies,[27,278] but not all.[32,283] One of those patterns, the estrogen-related dietary pattern (ERDP) developed by our group, was based on an estrogen profile that is specific to breast cancer risk: high unconjugated estradiol (E2) and a low ratio of 2- to 16-hydroxylated metabolites (2/16).

In the present analysis, we aimed to assess the relationship between a lifestyle score based on estrogen-related lifestyle factors and postmenopausal breast cancer risk. We created the estrogen-related lifestyle score (ERLS) using the ERDP, alcohol consumption, body mass index (BMI), and PA as scoring components, and examined associations with overall postmenopausal breast cancer and by estrogen receptor (ER) subtype, with consideration of potential effect modifiers. We hypothesized that higher ERLS scores, representative of a lower combined estrogenic effect of lifestyle factors, would be inversely associated with postmenopausal breast cancer, and that more substantial associations would be present for ER+ cases, and among strata of effect modifiers associated with lower estrogen exposure.

6.3 Methods

6.3.1 Study Population

The Prostate, Lung, Colorectal & Ovarian Cancer Screening Trial (PLCO) is an initiative of the National Cancer Institute to examine the effects of screening on cancer prognosis and mortality. Design and implementation have been described in detail elsewhere.[33] Briefly, recruitment of 76,685 men and 78,216 women aged 55 to 74 took place at 10 different screening centers across the United States between 1993 and 2001. Women in the screening arm participated in chest x-ray, flexible sigmoidoscopy, a digital rectal examination, a CA-125 blood test and transvaginal ultrasound. The current analyses used only data from women randomized to the intervention arm of the study (n=39,104) who completed a dietary questionnaire (DQX) at baseline, because participants in the control arm completed a different dietary questionnaire three years post-baseline. The study population was limited to women who completed the baseline questionnaire, had a valid DQX (between 1st and 99th percentiles of caloric intake, <8 missing line items), and without a personal history of cancer (except for non-melanoma skin cancer) at baseline, bringing the sample to 28,438. Participants were further excluded if they had an extreme (<15 or >55 kg/m^2; n=74) or missing (n=179) BMI, did not have data on PA (n=112), or if they did not contribute any person-time (n=58). After excluding participants with missing covariate data (n=862) the final analytic sample comprised 27,153 participants. A subsample of women had estrogen metabolite (EM) data, measured by liquid chromatography-tandem mass spectrometry assay of serum samples collected at baseline. This subsample, used in the development of the ERDP,[278]

came from a nested case-control study[35] and is comprised of 386 controls and 250 confirmed breast cancer cases who were diagnosed >2 years after blood sample donation.

6.3.2 Data Collection

Eligible participants filled out a questionnaire with information on demographics, medical history, family history, lifestyle factors, and recent history of participation in screening examinations at baseline. Participants self-reported their height and weight, which was used to calculate BMI. Dietary data were collected via the DQX, a 137-item food frequency questionnaire designed specifically for PLCO to assess typical frequency of intake over the past year. Nutrient and food intake amounts were calculated using US dietary data and the pyramid food group servings database from the US Department of Agriculture (USDA).[263] Separate line items were included for beer, liquor, and wine; which were used to calculate alcohol drinks per day. The DQX also contained a question on the number of hours per week spent performing vigorous PA, with the response categories of: <1, 1, 2, 3, and ≥4.

6.3.3 Calculating of ERLS Scoring

The dietary component of the ERLS was characterized using previously described ERDP scores (ref to submitted paper). Briefly, reduced rank regression modeling was performed to identify a dietary pattern that was associated with serum levels of unconjugated E2 and the 2/16 ratio in a nested case-control of 653 postmenopausal women. The newly developed ERDP is comprised of non-whole/refined grains, tomatoes, cruciferous vegetables, cheese, fish/shellfish high in ω-3 fatty acids, franks/luncheon

meats, nuts and seeds, other vegetables, fish/shellfish low in ω-3 fatty acids, yogurt, and coffee. Intakes of these food groups were centered and scaled, then multiplied by their corresponding model weights. The total ERDP score was calculated by summing over the weighted intakes. Higher ERDP scores are positively associated with unconjugated E2 and inversely associated with the 2/16 ratio. The dietary component of the ERLS score was based on the median ERDP score (-0.0206419) for the analytic PLCO population. Women with a score greater than or equal to the median received a 0, as those diets are hypothesized to be positively associated with estrogen metabolism and subsequent breast cancer risk. Women with an ERDP score below the median received a 1.

Scoring parameters for the remaining ERLS components are similar to those outlined in the WCRF/AICR Second Expert Report, and the USDA's 2015 Dietary Guidelines for Americans.[13,25] Due to the strength of evidence for associations between alcohol intake and obesity status with breast cancer risk, and evidence of an estrogenic effect, these variables were given a stronger weight in the scoring of the ERLS,[13] by using a three-level variable rather than two-level variable in the scoring. For alcohol intake, women who abstained from drinking (0 drink/week) were scored a 2; women who consumed >0 to 7 drinks/week were scored a 1; and those who consumed >7 drinks/week were scored a 0. Women were scored a 2 if they were normal weight (BMI <25.0 kg/m^2), a 1 if overweight (BMI 25.0-29.9 kg/m^2), and 0 if obese (BMI ≥30.0 kg/m^2). For PA, women who reported >2 hours/week of vigorous PA were considered active and scored a 1, and those who reported ≤2 hours/week were scored a 0. The score for each of the four different ERLS components was summed. Women with the minimum score of 0 were hypothesized to have the largest risk profile, and those with a maximum of 6 were

hypothesized to have the lowest combined risk profile from estrogen-related lifestyle factors. A summary of the ERLS scoring is portrayed in **Table 3.3**.

6.3.4 Breast Cancer Ascertainment

Incident breast cancer cases were identified primarily through self-report via annually mailed follow-up questionnaires. Follow-up was from start of enrollment in 1993 through December 31, 2009. Other sources of ascertainment included the National Death Index, physician reports, state cancer registries, and next of kin reports. Over 96% of the cases were confirmed through hospital records.[252] In the analytic cohort, a total of 1,568 incident breast cancer cases occurred. A supplemental form was implemented in 2007 to capture more detailed information about the diagnosis, including estrogen receptor status. Data on ER status was available for 70% of cases.

6.3.5 Statistical Approach

Baseline comparisons of participant characteristics by categories of the ERLS were performed using t-tests and chi-square tests for continuous and categorical variables, respectively. Pearson correlation coefficients were used to quantify the strength of the relationship between the ERLS and EMs in the subsample of women with data on EMs. Cox proportional hazards models were applied to analyze the relationship between the ERLS and incident breast cancer events, with person-time calculated from date of completed DQX to end of follow-up or event.[284] The proportional hazards assumption was evaluated using Martingale-based residuals and was not violated by exposure variables or covariates.[285] The ERLS was grouped as follows: ≤2 (referent group), 3, 4, or

≥5. The three lowest scores (0, 1, and 2) and the two highest scores (5 and 6) were combined into single categories due to low numbers of cases. The first category hypothetically represents lifestyles with a higher exposure to estrogen. The hazard ratio (HR) and 95% confidence intervals (CI) also were calculated for the continuous ERLS score variable, and the p-value reported as a test for trend. Demographic factors of age (years), race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian, other) and study center (10 categories) were included in the multivariable-adjusted models, along with total caloric intake (kcal/day) for their putative roles as confounders for breast cancer. The remaining covariates included in multivariable-adjusted models were chosen using stepwise model selection with entry/exit criteria of p=0.2. Further adjustment for hormone replacement therapy (HRT) use (current; former; never; unknown), family history of breast cancer (yes; no; unknown), education (less than high school; high school and some college; college degree; graduate degree), BMI at age 20 (kg/m^2), bilateral oophorectomy (yes; no), parity (6 categories), and age at menopause (5 categories) was included in the multivariable models. Age at first birth, age at menarche, oral contraceptive use, smoking status, and prior hysterectomy also were considered as potential confounders but were not included after performing the stepwise model selection. Effect modification by baseline HRT use (yes; no) and parity (nulliparous; parous) was examined in stratified results, and by including an interaction term in the model. All models were performed with overall breast cancer and within strata of ER subtype. A competing risk model was performed to assess a differential association for the ERLS on ER+ and ER- subtypes using a Wald test for heterogeneity in the stratified Lunn-McNeil approach.[259] In secondary analyses, we evaluated associations between

individual components of the ERLS and postmenopausal breast cancer with additional adjustment for each of the ERLS components that were not the main independent variable of interest. Additionally, to evaluate whether the observed association between the ERLS and postmenopausal breast cancer was primarily influenced by a single ERLS component, we removed the components one at a time from the total ERLS score to see if the estimate of association with breast cancer changed significantly. All statistical tests were two-sided at $\alpha=0.05$ and all analyses were conducted using SAS 9.4 (SAS Institute, Cary, NC).

6.4 Results

Over an average follow-up of 10.9 years, 1,576 incident cases of breast cancer were reported, with 1,261 of those cases being invasive. Among cases where ER status was ascertained, 1,089 were ER+ and 187 were ER-. In our subsample of participants with EM data that was used to derive the ERDP, the ERLS was moderately correlated with unconjugated E2 ($r=-0.33$; $p<0.01$) and the 2/16 ratio ($r=0.20$; $p<0.01$). The distribution of characteristics across ERLS categories for the full analytic cohort are shown in **Table 6.1**. Participants in the highest ERLS category had the lowest occurrence of total and ER+ breast cancer. In addition, they had the lowest total caloric intake, lowest proportion of non-Hispanic Whites but highest proportion of Asians, were more educated, and had the highest proportions of HRT users and never smokers at baseline.

In Cox models with varying levels of adjustment, participants in the highest ERLS category, representing lifestyles hypothesized to have the least estrogenic potential, experienced the lowest risk of postmenopausal breast cancer compared to the

lowest ERLS category (**Table 6.2**). In the multivariable-adjusted model, participants with an ERLS of 4 or ≥5 (lower estrogen) had a 23% (HR: 0.77; 95%CI: 0.67-0.89) and 34% (HR: 0.66; 95%CI: 0.56-0.78) reduction in risk of breast cancer, respectively, compared to those with an ERLS ≤2 (higher estrogen) (p-trend<0.0001). A 1-unit increase in ERLS was associated with a 11% lower risk (HR: 0.89; 95%CI: 0.85-0.92) after adjustment. Estimates were similar for invasive cases only. When restricting to ER+ subtype, the magnitude of the inverse associations strengthened slightly for those with an ERLS of 4 (HR: 0.73, 95%: 0.62-0.87) and ERLS ≥5 (HR: 0.63; 95%CI: 0.51-0.77). No significant effect estimates were observed for ER- subtypes, but the HR for ERLS ≥5 was reduced and results from the competing risk model indicated there was no differential association for the different ER subtypes (p=0.62). There was no evidence of effect modification by baseline HRT use ($p_{interaction}$=0.54) or parity ($p_{interaction}$=0.75) (**Table 6.3)**.

Table 6.4 shows results from investigations of individual ERLS scoring components. In all models, the category that was associated with a score of 0, representative of higher estrogen exposure, was the referent. A modest reduction in risk was observed in participants with score of 1 for the ERDP (HR: 0.92; 95%CI: 0.83-1.02). Significant reductions in risk were seen among individuals with an alcohol score of 2 (HR: 0.79; 95%CI: 0.66-0.95), or 1 (HR: 0.83; 95%CI: 0.72-0.95); individuals with a BMI score of 2 (HR:0.72; 95%CI: 0.62-0.83), or 1 (HR: 0.88; 95%CI: 0.76-1.00); and for those with a score of 1 for PA (HR: 0.92; 95%CI: 0.83-1.02). The estimates of association for the ERLS remained relatively unchanged after removing individual components, one at a time (**Table 6.5**).

6.5 Discussion

In this large prospective cohort study of postmenopausal women, our findings suggest that the combined effect of modifiable lifestyle factors, namely diet, alcohol intake, BMI, and PA, is associated with risk of postmenopausal breast cancer. Specifically, women who were consuming a diet with less estrogenic potential, less alcohol, had a lower BMI, and were engaging in more physical activity were at reduced risk for breast cancer compared to women with less healthy lifestyles. A 1-unit increase in the ERLS score towards the direction of a lifestyle that was hypothesized to have lower estrogen exposure was associated with a 11% reduction in risk. The ERLS was moderately correlated with two EMs thought to be important indicators of breast cancer risk in a subsample of women. However, the association between ERLS and breast cancer did not differ by ER subtypes. The association was not modified by HRT use or parity.

Considering the prominence of an estrogenic influence on the development of breast cancer, the ERLS was developed to characterize the combined effect of estrogen-related lifestyle factors. Other lifestyle components, such as smoking or breastfeeding, were omitted from the ERLS as evidence of an estrogenic disease mechanism is not substantial.[67] All individual components of the ERLS exhibited inverse associations with postmenopausal breast cancer in multivariable-adjusted models, but none of the estimates of association were larger than their combined effect in the ERLS. According to the World Cancer Research Fund (WCRF) and American Institute for Cancer Research's (AICR) 2017 Continuous Update Project (CUP),[24] there is strong evince of increasing risk of postmenopausal breast cancer with body fatness (represented by BMI in the

ERLS) and alcohol. The CUP also has designated PA to have strong evidence of an influence on breast cancer risk, therefore we anticipated seeing an association.[24] Furthermore, in a prior PLCO investigation, a 1-unit increase in ERDP scores was associated with a significant 9% increase in risk of developing postmenopausal breast cancer.[278] The association between ERLS and postmenopausal breast cancer remained significant, with relatively no attenuation in effect estimates, even after individual ERLS components were removed from the total score. These results suggest there was no single component of the ERLS that drove the observed significant association in models with total ERLS.

To the best of our knowledge, this was the first application of a lifestyle score with a focus on estrogen metabolism as the primary mechanistic pathway. Prior research on lifestyle scores and breast cancer in prospective studies have yielded similar results. An *a priori* healthy lifestyle index score (HLIS) based on diet, tobacco use, alcohol, PA, and BMI reported 21% lower risk of breast cancer (HR: 0.79, 95% CI: 0.73-0.85) among the most healthy group in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort.[246] When the HLIS was applied in the same cohort, but with a slight dietary modification to include fish, folate, glycemic index, and other breast cancer risk-specific dietary components, the estimate of the inverse association was slightly stronger (HR: 0.74; 95% CI: 0.66–0.83).[16] The association was strongest for ER-/progesterone receptor (PR)- breast cancer (HR: 0.60; 95% CI: 0.40-0.90), but also significant for ER+/PR+ (HR: 0.81; 95% CI: 0.67-0.98), suggesting disease pathways that did not influence estrogen may have played a role.[16] Also using data from EPIC, a lifestyle score was developed to evaluate adherence to the WCRF/AICR

recommendations on body fatness, PA, energy dense foods and drinks, plant foods, animal foods, alcohol use, and breastfeeding in women. Compared to the lowest scores, all categories showed a significant inverse association with breast cancer, with the strongest association in those with greatest adherence to the prevention guidelines (HR: 0.84; 95% CI: 0.78-0.90).[14] Adherence to WCRF/AICR recommendations has exhibited positive associations with breast cancer risk in other populations, as well, [247–249] including the Iowa Women's Health Study where association did not differ in the presence of non-modifiable risk factors for breast cancer.[249]

Evidence from case-control studies have shown similar, yet stronger associations. In a case-control study of Mexican women, those in the highest quintile of a healthy index comprised of diet, PA, alcohol consumption, and tobacco smoking had 80% less odds of developing postmenopausal breast cancer compared to the lowest quintile (odds ratio (OR): 0.20; 95% CI: 0.11-0.37).[15] Increasing scores associated with a lifestyle score focused on limiting red meat, cream, and cheese; consuming more white meat, fish, fruit and vegetables; lower alcohol consumption; not smoking; higher PA; lower BMI; and longer cumulative duration of breastfeeding was associated with a reduction in risk among indigenous New Zealanders (OR: 0.47; 95% CI: 0.23-0.94), but not among non-indigenous participants (OR: 0.86; 95% CI: 0.67-1.11), when comparing the highest to lowest quartiles.[250]

There is evidence that high levels of circulating unconjugated E2 and a low 2/16 ratio may be representative of an estrogen profile that increases the risk of postmenopausal breast cancer.[11] In our subsample of women with EM data, the ERLS was inversely and positively correlated with unconjugated E2 and the 2/16 ratio,

respectively. Additionally, each component of the ERLS has been associated with estrogen metabolism.[140,212,278] Therefore, it is plausible that the combined effect of these lifestyle behaviors on postmenopausal breast cancer risk works through an influence on estrogen metabolism. Dietary behaviors are known to influence the intestinal microbiota,[266] which can subsequently influence excretion or reabsorption of active estrogens.[267] Alcohol consumption may increase aromatase activity, promoting the conversion of testosterone into estrogen.[286] Adipose tissue is the largest source of endogenous estrogen in postmenopausal women,[2] and there is strong evidence for a positive linear association between adipose tissue and estrogen levels in postmenopausal women.[207] The inverse association between PA and estrogen may be a result of reducing adipose-derived estrogen, or possibly through increased levels of SHBG, limiting the amount of available estrogen in active tissues.[24,62]

Some limitations should be considered. Similar to most prospective investigations, there is the potential for bias due to the selection of subjects, loss to follow-up, and measurement error. Although food frequency questionnaires may not generate accurate estimates for absolute intakes of nutrients, they are useful for ranking individuals, and only food or food groups (not nutrients) intakes were utilized in this study.[173] The use of BMI is an imperfect proxy for adiposity, and BMI values were derived from self-reported height and weight. However a validation study in a similar U.S. population showed strong correlation between self-reported and measured weight.[287] Our ability to detect an association for ER- cases was limited due to low numbers, however, this was not an issue for ER+ cases. A limitation for the PLCO study population is the lack of racial/ethnic diversity. However, non-Hispanic White women

experience the highest incidence of breast cancer in the US, so results are generalizable to this high-risk group.

There are many strengths to our analysis, as well. The use of a large, prospective cancer cohort provided adequate power to detect small associations with complete information on known risk factors to appropriately adjust for confounders. The inclusion of the ERDP and pre-identification of a plausible mechanistic pathway aided in making a meaningful interpretation of our results. This was a novel approach to developing a lifestyle score that is disease- and mechanism-specific.

In conclusion, our findings suggest that modifiable lifestyle behaviors have a combined effect on postmenopausal breast cancer risk, possibly through an alteration of estrogen metabolism. A lifestyle that is characterized by consumption of a diet with low estrogenic potential, low alcohol consumption, a low BMI, and high levels of PA may help to lower the risk of developing breast cancer in postmenopausal women. A collective change in lifestyle is likely more influential than focusing on specific behaviors.

6.6 Tables

Table 6.1 Population characteristics across estrogen related lifestyle score (ERLS) categories

	ERLS			
	≤2	3	4	≥5
n	7,469	7,565	7,345	4,774
Breast cancer cases				
Total	459	491	400	226
Invasive	368	401	308	184
ER+	321	342	272	154
ER-	56	56	48	27
ERDP score (mean ± SD)	0.31 ± 0.63	0.05 ± 0.62	-0.14 ± 0.58	-0.39 ± 54
(ERLS: 0) ≥ median, %	78.6	56.8	38.6	12.0
(ERLS: 1) < median, %	21.4	43.2	61.4	88.0
Alcohol (drinks/week, mean ± SD)	4.1 ± 8.1	3.4 ± 7.1	2.3 ± 5.1	0.8 ± 1.5
(ERLS: 0) >7, %	22.2	17.2	9.5	0.0
(ERLS: 1) >0-7, %	69.4	65.6	69.1	53.1
(ERLS: 2) 0, %	8.4	17.2	21.4	46.9
BMI (kg/m^2, mean ± SD)	31.9 ± 5.5	27.3 ± 4.7	24.6 ± 3.3	23.0 ± 2.3
(ERLS: 0) ≥30, %	64.3	20.5	3.6	0.0
(ERLS: 1) 25.0-29.9, %	31.5	50.9	37.6	11.6
(ERLS: 2) <25, %	4.2	28.6	58.8	88.4
Hours of vigorous PA per week (%)				
(ERLS: 0) ≤2	81.1	51.4	28.6	8.8
(ERLS: 1) >2	18.9	48.6	71.4	91.2
Age (mean ± SD)	61.9 ± 5.2	62.4 ± 5.3	62.7 ± 5.4	63.0 ± 5.5
Total caloric intake (kcal/day, mean± SD)	1,894 ± 639	1,763 ± 598	1,677 ± 572	1,577 ± 529
BMI at age 20 (kg/m^2, mean ± SD)	22.3 ± 3.3	21.3 ± 2.8	20.7 ± 2.4	20.4 ± 2.1

HRT status (%)					
	Current	45.4	51.7	54.3	57.2
	Former	17.8	15.9	15.4	14.8
	Never	36.2	32.1	29.8	27.7
	Unknown	0.6	0.3	0.5	0.3
Race (%)					
	White, Non-Hispanic	92.3	91.8	91.4	89.3
	Black, Non-Hispanic	5.2	4.7	3.3	2.4
	Hispanic	1.2	1.2	1.2	1.1
	Asian	0.6	1.8	3.5	6.8
	Other	0.7	0.5	0.6	0.4
Smoking (%)					
	Current	9.5	9.8	8.2	6.7
	Former	37.7	34.5	33.3	28.3
	Never	52.8	55.7	58.5	65.0
Education (%)					
	< HS	6.7	6.1	4.7	4.6
	HS grad and some college	68.2	64.0	62.3	61.7
	College grad	13.5	15.8	17.4	16.6
	Postgraduate	11.6	14.1	15.6	17.1
Live births (%)					
	None	6.6	7.6	7.6	7.8
	1	6.9	7.3	7.1	7.3
	2	21.7	23.3	25.5	26.3
	3	25.7	25.3	26.1	26.2
	≥ 4	39.1	36.5	33.7	32.4
Age at menopause (%)					
	< 40	15.3	13.5	12.6	12.9
	40-44	14.1	14.0	13.4	14.3

45-59	22.7	23.1	23.4	25.3
50-54	36.4	37.8	38.9	36.8
≥55	11.5	11.6	11.7	10.7
Bilateral oophorectomy (%)				
No	88.0	88.9	89.7	89.2
Yes	12.0	11.1	10.3	10.8
Family history of breast cancer (%)				
No	84.3	84.7	85.4	85.2
Yes	14.5	14.4	13.7	14.0
Unknown	1.2	1.0	0.9	0.9

BMI: body mass index; ER: estrogen receptor; ERDP: estrogen related dietary pattern; ERLS: estrogen related lifestyle score; HRT: hormone replacement therapy; HS: high school; PA: physical activity; SD: standard deviation

Table 6.2 Hazard ratios (95% CI) for the relationship between estrogen related lifestyle score (ERLS) and postmenopausal breast cancer

	ERLS				Estimate for continuous ERLS,[a] p-trend
	≤2	3	4	≥5	
Total breast cancer					
No. of cases	459	491	400	226	
Age-adjusted	1.00 (ref)	1.05 (0.92, 1.19)	0.87 (0.76, 0.99)	0.75 (0.64, 0.88)	0.92 (0.89, 0.96) p<0.0001
Age- and HRT-adjusted	1.00 (ref)	1.02 (0.90, 1.16)	0.84 (0.73, 0.96)	0.72 (0.61, 0.84)	0.91 (0.88, 0.95) p<0.0001
Multivariable-adjusted[b]	1.00 (ref)	0.97 (0.85, 1.11)	0.77 (0.67, 0.89)	0.66 (0.56, 0.78)	0.89 (0.85, 0.92) p<0.0001
Invasive					
No. of cases	368	401	308	184	
Age-adjusted	1.00 (ref)	1.07 (0.92, 1.23)	0.83 (0.72, 0.97)	0.76 (0.64, 0.91)	0.92 (0.88, 0.96) p=0.0003
Age- and HRT-adjusted	1.00 (ref)	1.05 (0.91, 1.21)	0.81 (0.69, 0.94)	0.73 (0.61, 0.88)	0.91 (0.87, 0.95) p<0.0001
Multivariable-adjusted[b]	1.00 (ref)	0.99 (0.86, 1.15)	0.74 (0.63, 0.87)	0.67 (0.56, 0.82)	0.89 (0.85, 0.93) p<0.0001
ER+					
No. of cases	321	342	272	154	
Age-adjusted	1.00 (ref)	1.04 (0.89, 1.21)	0.84 (0.71, 0.98)	0.73 (0.60, 0.88)	0.91 (0.87, 0.96) p=0.0001
Age- and HRT-adjusted	1.00 (ref)	1.01 (0.87, 1.18)	0.81 (0.69, 0.95)	0.70 (0.57, 0.84)	0.90 (0.86, 0.94) p<0.0001
Multivariable-adjusted[b]	1.00 (ref)	0.96 (0.82, 1.12)	0.73 (0.62, 0.87)	0.63 (0.51, 0.77)	0.87 (0.83, 0.92) p<0.0001
ER-					
No. of cases	56	56	48	27	
Age-adjusted	1.00 (ref)	1.04 (0.72, 1.52)	0.91 (0.61, 1.34)	0.79 (0.50, 1.26)	0.95 (0.84, 1.06) p=0.34

Age- and TEI-adjusted	1.00 (ref)	1.04 (0.71, 1.51)	0.90 (0.61, 1.34)	0.78 (0.49, 1.25)	0.94 (0.84, 1.06) p=0.32
Multivariable-adjusted[b]	1.00 (ref)	1.04 (0.71, 1.53)	0.92 (0.61, 1.38)	0.84 (0.52, 1.37)	0.96 (0.85, 1.09) p=0.52

CI: confidence interval; ER: estrogen receptor; ERLS: estrogen related lifestyle score; HRT: hormone replacement therapy; TEI: total energy intake
aHR corresponds to 1-unit increase in ERLS score.
bIncludes adjustment for age, TEI, HRT, education, BMI at age 20, bilateral oophorectomy, parity, age at menopause, family history of breast cancer, race/ethnicity, and study center.

Table 6.3 Hazard ratios (95% CI) for the relationship between the estrogen related lifestyle score (ERLS) and postmenopausal breast cancer within strata of estrogen-related risk factors[a]

		ERLS				
		≤2	3	4	≥5	p interaction[b]
HRT use at baseline						0.54
	No	1.00 (ref)	0.97 (0.80, 1.17)	0.71 (0.57, 0.88)	0.67 (0.51, 0.87)	
	Yes	1.00 (ref)	0.97 (0.82, 1.16)	0.82 (0.68, 0.98)	0.66 (0.53, 0.82)	
Parity						0.75
	Nulliparous	1.00 (ref)	0.82 (0.52, 1.29)	0.67 (0.41, 1.08)	0.49 (0.27, 0.90)	
	Parous	1.00 (ref)	0.99 (0.86, 1.14)	0.78 (0.67, 0.90)	0.67 (0.56, 0.80)	

CI: confidence interval; ERLS: estrogen related lifestyle score; HRT: hormone replacement therapy; TEI: total energy intake

[a]Includes adjustment for age, TEI, HRT, education, BMI at age 20, bilateral oophorectomy, parity, age at menopause, family history of breast cancer, race/ethnicity, and study center.

[b]P-value for the product term of ERDP quartiles with the potential effect modifier.

Table 6.4 Hazard ratios (95% CI) for the relationship between the individual estrogen related lifestyle score (ERLS) components and postmenopausal breast cancer

	No. of cases	Age-adjusted	Age- and HRT-adjusted	Multivariable-adjusted[a]
ERDP score				
≥ median	827	1.00 (ref)	1.00 (ref)	1.00 (ref)
< median	749	0.91 (0.82, 1.00)	0.90 (0.82, 1.00)	0.92 (0.83, 1.02)
Alcohol (drinks/week)				
>7	264	1.00 (ref)	1.00 (ref)	1.00 (ref)
>0 to 7	1,025	0.79 (0.69, 0.91)	0.80 (0.70, 0.92)	0.83 (0.72, 0.95)
0	287	0.71 (0.60, 0.84)	0.71 (0.60, 0.84)	0.79 (0.66, 0.95)
BMI (kg/m^2)				
≥30	388	1.00 (ref)	1.00 (ref)	1.00 (ref)
25.0 to 29.9	578	1.00 (0.88, 1.13)	0.97 (0.85, 1.10)	0.88 (0.76, 1.00)
25	610	0.90 (0.80, 1.03)	0.86 (0.76, 0.98)	0.72 (0.62, 0.83)
Hours of vigorous PA per week				
≤2	733	1.00 (ref)	1.00 (ref)	1.00 (ref)
>2	843	0.95 (0.86, 1.05)	0.93 (0.84, 1.03)	0.92 (0.83, 1.02)

BMI: body mass index; CI: confidence interval; ERDP: estrogen related dietary pattern; ERLS: estrogen related lifestyle score; HRT: hormone replacement therapy; PA: physical activity; TEI: total energy intake

aIncludes adjustment for each other ERLS component that is not the main predictor, age, TEI, HRT, education, BMI at age 20, bilateral oophorectomy, parity, age at menopause, family history of breast cancer, race/ethnicity, and study center.

Table 6.5 Hazard ratios (95% CI) for the relationship between estrogen related lifestyle score (ERLS) and postmenopausal breast cancer removing individual ERLS components from the total score[a]

Component removed from total ERLS:	Estimate for continuous ERLS,[b] p-trend
ERDP	0.88 (0.83, 0.92) p<0.0001
Alcohol	0.89 (0.85, 0.93) p<0.0001
BMI	0.90 (0.85, 0.95) p=0.0003
PA	0.87 (0.83, 0.92) p<0.0001

[a]Includes additional adjustment for age, TEI, HRT, education, BMI at age 20, bilateral oophorectomy, parity, age at menopause, family history of breast cancer, race/ethnicity, and study center.
[b]HR corresponds to 1-unit increase in ERLS score.

CHAPTER 7

DIET, LIFESTYLE, AND ESTROGEN METABOLISM IN RELATION TO POSTMENOPAUSAL BREAST CANCER: A SYNTHESIS OF DISSERTATION FINDINGS

7.1 Summary of findings

We characterized women's diets based on associations with an estrogen profile that is hypothesized to be associated with increased postmenopausal breast cancer risk; high unconjugated E2 and a low 2/16 ratio. Starting with 32 food and beverage groups, we identified 11 key contributors to the variation in the EMs of interest. Intakes of non-whole/refined grains, tomatoes, cruciferous vegetables, cheese, fish/shellfish high in ω-3 fatty acids, franks/luncheon meats were positively weighted for ERDP scoring; whereas intakes for nuts and seeds, other vegetables, fish/shellfish low in ω-3 fatty acids, yogurt, coffee were negatively weighted. Next, the ERDP was scored in two prospective cohorts of postmenopausal women and examined for an association with breast cancer risk. In the PLCO, the cohort from which the ERDP was developed, a positive association between the ERDP and postmenopausal breast cancer risk was observed, in that the highest quartile of the ERDP was associated with a 20% increased risk of invasive breast cancer compared to the lowest quartile. However, the ERDP was not associated with postmenopausal breast cancer risk among women in the SS.

Possible reasons for the different findings between the two study populations could be because all SS participants had a family history of breast cancer which reflected shared genetic and early life environment, differences in dietary measurements and subsequent ERDP scoring distributions between the studies, or due to a chance finding in the PLCO when in truth there is no association. To examine the potential role of inherited risk affecting the observed associations in the PLCO, we conducted further analyses stratified by family history of breast cancer. As shown in **Table 7.1,** no association was observed when restricting to PLCO participants with a family history of breast cancer, as is characteristic of the full SS cohort, thus supporting the idea that high inherited risk from shared genetic profiles and early life environments may be masking an association with diet in the SS. **Table 7.2** shows how intakes of some food groups differ significantly between the two populations, likely due to differences in the descriptions and number of line items containing foods within those groups in the different FFQs. These differences in intake measurement translated to different distributions of ERDP scores across the two populations. Scores in PLCO were slightly skewed right, whereas scores in SS were slightly skewed left. Considering the more negative distribution in SS, it is possible that participants in SS did not consume enough pro-estrogenic foods to observe an association. Lastly, participants in SS had an average shorter duration of follow-up. Therefore, if their dietary estrogenic potential measured at baseline requires a longer time period to influence breast cancer risk, we may have only been able to evaluate the association effectively in PLCO. This would not be relevant if their diet at baseline is similar to their diet in previous years, but literature on the stability of dietary patterns in adulthood is varied, so it is difficult to defend that assumption.[288–291]

Lastly, the ERDP was incorporated into an estrogen-related lifestyle score (ERLS) with other estrogen-related lifestyle factors. The ERLS was comprised of the ERDP, alcohol consumption, BMI, and PA; with increasing scores hypothesized to have a combined anti-estrogenic potential. An inverse association between the ERLS and postmenopausal breast cancer risk was observed in PLCO, with women conforming to more of the healthy lifestyle factors having a 34% reduced risk of breast cancer compared to fewer healthy lifestyle factors, supporting the hypothesis that modifiable lifestyle factors related to lower estrogenic potential are associated with reduced risk of breast cancer.

7.2 Biological mechanisms

The ERDP was hypothesized to be related to postmenopausal breast cancer risk through a biologic mechanism related to estrogen metabolism. One possible way by which ERDP food groups may affect estrogen metabolism is through an influence on microbiome diversity.[265] Diversity of the intestinal microbiome, which is strongly influenced by dietary behaviors, can impact many important physiological processes, such as whether or not estrogens are excreted through feces or transformed to their unconjugated forms and subsequently reabsorbed.[267] If reabsorbed, there is a greater estrogenic exposure throughout the body. Similarly, there is evidence of microbial effects on interconversions of the parent estrogens and hydroxylation down the 16-pathway from *in vitro* and human studies, suggesting the microbiome may also influence estrogen metabolism.[269,270] The intestinal microbiome is strongly influenced by fiber intake, or lack thereof, through consumption of grains and vegetables, both of which are included in

the ERDP.[266] The ERDP also is comprised of animal products, such as meats, cheese, and yogurt, which can impact microbiome diversity.[271–273] Therefore, diet may influence breast cancer risk through an influences on a woman's estrogen profile, mediated by microbial effects. Many of the foods in the ERDP are also characteristic of a Western dietary pattern , which has been associated with systemic inflammation.[275] Conversely, coffee, also a part of the ERDP, has exhibited anti-inflammatory effects.[274] Inflammation may play a role in mammary tumor development, therefore it may also play a role in a potential association between the ERDP and breast cancer.[276]

In addition to the aforementioned mechanistic pathways for the relationship between the ERDP and postmenopausal breast cancer, it is possible other components of the ERLS also work through estrogen metabolism. Alcohol consumption may increase aromatase activity, promoting the conversion of testosterone into estrogen.[286] Adipose tissue is the largest source of endogenous estrogen in postmenopausal women,[2] and there is strong evidence for a positive linear association between adipose tissue and estrogen levels in postmenopausal women.[207] The inverse association between PA and estrogen may be a result of reducing adipose-derived estrogen, or possibly through increased levels of SHBG, limiting the amount of available estrogen in active tissues.[24,62] Similarly to the ERDP, other components of the ERLS may work through inflammatory mechanisms, as well, with alcohol, adipose tissue, and physical activity all exhibiting associations with inflammatory markers.[292]

7.3 Implications for public health

The results of the present dissertation help to address a critical gap in translational breast cancer research. The burden of breast cancer is extensive, as it accounts for nearly a third of all cancers diagnosed among women.[2] Costs of treatment are extensive, as is the potential for secondary health effects among the large numbers of breast cancer survivors. To help ease this burden, primary prevention methods utilizing modifiable lifestyle factors are needed.[43] Diet is a commonly investigated lifestyle factor, though previous studies have largely yielded inconsistent results in relation to breast cancer. By focusing on the whole diet and a biologic mechanism specific to breast cancer, this dissertation adds to the literature on whether disease-specific dietary recommendations are warranted. Our findings suggest that a diet associated with estrogen metabolism may influence breast cancer risk, although the dietary pattern established in one study was not associated with breast cancer in a different study population. Thus, further research capturing the "optimal" diet for estrogen metabolism across multiple populations is warranted. Furthermore, we have shown that the combined effect of adopting lifestyle factors associated with lower estrogen exposure may be efficacious to reduce the risk of postmenopausal breast cancer.

Overall, our results suggest that a diet low in non-whole/refined grains, tomatoes, cheese, franks/luncheon meats; while high in nuts and seeds, cruciferous vegetables, other vegetables, fish/shellfish, yogurt, and coffee may protect against breast cancer through an influence on estrogen metabolism. More research is needed to determine the effects of an estrogen-related diet in other populations, including those with a strong inherited risk or different dietary habits. In addition to diet, a lifestyle that is

characteristic of a healthy body weight, low consumption of alcohol, and increased PA may help to prevent postmenopausal breast cancer incidence.

7.4 Strengths and limitations

Some limitations must be considered when interpreting results from the present dissertation. There was the potential for bias due to the selection of subjects, such as in SS participants where some women with a particularly strong inherited risk may have taken necessary lifestyle changes to reduce their risk. However, all women in the cohort are aware of their family history of breast cancer therefore we do not expect a differential effect based on inherited risk. Loss to follow-up represents the potential for selection bias, however response rates in SS have been >90% for all survey periods. In PLCO, over 75% of participants were followed for at least 10 years and 95% of participants were followed for at least 4 years, suggesting a low number of early drop outs. FFQs are prone to measurement error, and specifically have been shown to poorly estimate current total energy intake in relation to a recovery biomarker (doubly labeled water).[293,294] We used only food or food-group data from the FFQs, rather than macronutrient or micronutrient data, thus, eliminating one source of error that results from converting food intake to nutrient intake using food composition databases. Any dietary measurement error would likely have been non-differential with respect to breast cancer outcome, thus biasing effect estimates toward the null. Use of multiple FFQs to assess diet in adulthood would have reduced intra-individual variation and better captured the estrogenic potential of diet, however evidence of changes in adulthood diet is limited[288–291] and FFQs are designed to assess usual diet. Because FFQs contain a predetermined list of foods and

beverages, the use of FFQs to develop the ERDP limited our ability to identify all foods that were associated with EM in the RRR modeling, as compared with an open-ended dietary assessment method such as 24 hour recalls or food records. While the PLCO and SS populations were similar with regard to SES and race/ethnicity distributions, differences in FFQs across the two study populations may have limited the comparability of our findings from Aim #1 and #2. Differences in the descriptions and numbers of line items for certain food groups may have affected our observed associations through an impact on the distribution of ERDP scores. A minor limitation in regards to study populations was the lack of heterogeneity of race and ethnicities. However, our populations are predominately non-Hispanic White women who experience the highest incidence of breast cancer, so the results have major public health relevance. While we adjusted for important potential confounders, residual or unmeasured confounding cannot be ruled out. A low percentage of variation in EMs was explained by the ERDP, therefore it is difficult to assess the role of estrogen metabolism in explaining the association with breast cancer. However, the percentage of EM variation explained in the ERDP was similar and slightly larger than other RRR analyses using intermediate biomarkers.[32,295]

There are major strengths in the approach and design to offset some of the limitations. Analyses were conducted with information on known confounders and with enough power to detect moderately small effects through the use of large, prospective cancer cohorts. Follow up was substantial enough for an adequate number of events to accrue, although shorter duration in SS compared to PLCO may have contributed to the difference in results across the two study populations. Equally small proportions (~5%) of cases and controls were excluded because of missing exposure or covariate data for

both populations, therefore we do not expect missing data to have influenced the differential results.

Using RRR to create the ERDP based on EM biomarkers allowed for consideration of an *a priori* mechanistic hypothesis to facilitate interpretation of results. The EMs used to develop the ERDP were measured using a sensitive assay and have been shown to be strongly related to breast cancer risk in the PLCO population.[35] Finally, the use of prospective cohort studies where diet was assessed prior to disease diagnosis minimizes the potential for recall bias which can afflict case-control studies.

7.5 Suggestions for future work

Most large prospective cohort studies in the US and worldwide have used FFQs to assess usual diet. FFQs are generally less expensive and more feasible in large population-based studies than other dietary assessment methods, such as 24 hour recalls or foods records. However, the previously mentioned limitations resulting from using an FFQ to derive the dietary pattern may be improved upon through use of an open-ended dietary assessment tool, such as a 24-hour recall of food record. In doing so, all foods consumed that may influence estrogen metabolism in a given population can be measured. Similarly, the development of an estrogenic dietary pattern should be conducted in multiple populations with different dietary habits in order to examine how the contributing foods vary, or if associations with breast cancer vary depending on the diets of each population. Another suggestion for improvement in future studies is to measure EMs at multiple times to reduce intra-individual variability in the intermediate outcome used to develop the ERDP. To confirm a mechanistic pathway that works

through estrogen metabolism, construct validation needs to be performed, such as by evaluating the relationship between our diet score and serum EM in another study population. Studies with serum EMs measured at an intermediate time point between dietary exposure and breast cancer outcomes would help to clarify the potential role of estrogen metabolism. Alternatively, clinical trials could be effective in determining differences in EM levels across experimental groups of high and low adherers to the ERDP.

7.6 Concluding remarks

In conclusion, we developed a dietary pattern associated with a high-risk estrogen profile (high E2 and low 2/16 ratio) that is hypothesized to increase breast cancer risk. Women who had high ERDP scores tended to consume higher amounts of non-whole/refined grains, tomatoes, cheese, franks/luncheon meats; and lower amounts of nuts and seeds, cruciferous vegetables, other vegetables, fish/shellfish, yogurt, and coffee. A subsequent prospective investigation indicated that this estrogenic diet was associated with an increased risk of postmenopausal breast cancer risk in the study cohort in which it was developed, PLCO. However, application of the dietary pattern in a second population with a high inherited risk, SS, resulted in no association with breast cancer. Taking the results from Aims #1 and #2 together, we emphasize the importance of considering dietary assessment tools when comparing interpretations from *a posteriori* patterns across populations, as well as the need for studies of lifestyle factors across strata of participants with or without a family history of breast cancer.

When the dietary pattern was incorporated into a lifestyle score with alcohol intake, BMI, and PA, a combined effect on postmenopausal breast cancer risk was observed in the PLCO. A lifestyle that is characterized by consumption of a diet with low estrogenic potential, low alcohol consumption, normal-weight BMI, and high levels of PA may help to lower the risk of developing breast cancer in postmenopausal women. A collective change in lifestyle is likely more influential than focusing on specific behaviors.

7.7 Acknowledgements

This dissertation research was supported in part by Grant Number T32-GM081740 from NIH-NIGMS. Mr. Guinter also was supported by the Support to Promote the Advancement of Research and Creativity graduate research grant from the University of South Carolina's Office of the Vice President for Research. Dr. Sandler is supported by the Intramural Program of the NIH, NIEHS (Z01 ES-044005l). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

7.8 Tables

Table 7.1 Hazard ratios (95% CI) for the relationship between the estrogen related dietary pattern (ERDP) score and postmenopausal breast cancer in PLCO participants stratified by family history of breast cancer[a]

	ERDP quartiles				Estimate for continuous ERDP score[b]
	1st	2nd	3rd	4th	
Without family history of breast cancer					
Total breast cancer cases	297	328	331	356	
	1.00 (ref)	1.11 (0.95, 1.30)	1.11 (0.95, 1.30)	1.17 (0.99, 1.37)	1.11 (1.02, 1.21) p=0.02
Invasive cases	229	261	270	290	
	1.00 (ref)	1.15 (0.96, 1.37)	1.18 (0.99, 1.41)	1.23 (1.03, 1.48)	1.17 (1.06, 1.29) p=0.002
With family history of breast cancer					
Total breast cancer cases	69	64	72	75	
	1.00 (ref)	1.00 (0.71, 1.42)	1.05 (0.75, 1.47)	1.03 (0.73, 1.45)	1.01 (0.83, 1.22) p=0.94
Invasive cases	51	48	61	58	
	1.00 (ref)	0.99 (0.67, 1.48)	1.19 (0.82, 1.73)	1.10 (0.74, 1.63)	1.01 (0.81, 1.26) p=0.93

ERDP: estrogen-related dietary pattern

[a]Includes adjustment for age, TEI, BMI, BMI at age 20, HRT, alcohol use, education, bilateral oophorectomy, parity, age at menopause, PA, race/ethnicity, and recruitment center.

[b]HR corresponds to 1-unit increase in ERDP score.

Table 7.2 Comparison of food group intakes in PLCO and the Sister Study

	PLCO				Sister Study			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Total ERDP Score	-0.01	0.65	-4.51	6.58	-0.05	0.71	-8.23	4.67
Non-whole/refined grains (oz/day)	4.19	1.95	0.3	16.3	2.87	1.62	0.0	18.7
Tomatoes (cups/day)	0.43	0.30	0.0	8.5	0.27	0.21	0.0	3.1
Other vegetables (cups/day)	0.99	0.55	0.0	6.2	0.48	0.38	0.0	5.4
Cruciferous vegetables (cups/day)	0.28	0.26	0.0	3.9	0.23	0.30	0.0	4.7
Cheese (cups/day)	0.35	0.30	0.0	4.3	0.39	0.32	0.0	2.6
Yogurt (cups/day)	0.12	0.20	0.0	2.2	0.12	0.19	0.0	1.8
Fish/shellfish high in ω-3 fatty acids (oz/day)	0.16	0.19	0.0	2.9	0.15	0.20	0.0	3.8
Fish/shellfish low in ω-3 fatty acids (oz/day)	0.50	0.47	0.0	10.3	0.45	0.46	0.0	8.9
Franks and luncheon meats (oz/day)	0.23	0.30	0.0	6.4	0.52	0.45	0.0	5.8
Nuts and seeds (oz/day)	0.42	0.63	0.0	9.9	1.44	1.60	0.0	23.2
Coffee (cups/day)	2.48	3.03	0.0	17.4	1.50	1.48	0.0	9.0

PLCO: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; SD: standard deviation

REFERENCES

1. Stewart BW, Wild CP. *World Cancer Report 2014*.; 2014.

2. American Cancer Society. *Breast Cancer Facts & Figures 2015-2016*. Atlanta; 2015.

3. Sarrel PM. Androgen deficiency: menopause and estrogen-related factors. *Fertil Steril*. 2002;77 Suppl 4:S63-7. http://www.ncbi.nlm.nih.gov/pubmed/12007905. Accessed September 26, 2016.

4. Samavat H, Kurzer MS. Estrogen metabolism and breast cancer. *Cancer Lett*. 2015;356(2):231-243. doi:10.1016/j.canlet.2014.04.018.

5. MacMahon B. Reproduction and cancer of the breast. *Cancer*. 1993;71(10):3185-3188. doi:10.1002/1097-0142(19930515)71:10<3185::AID-CNCR2820711047>3.0.CO;2-C.

6. McPherson K, Steel CM, Dixon JM. ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *BMJ*. 2000;321(7261):624-628. http://www.ncbi.nlm.nih.gov/pubmed/10977847. Accessed September 27, 2016.

7. Brown SB, Hankinson SE. Endogenous estrogens and the risk of breast, endometrial, and ovarian cancers. *Steroids*. 2015;99:8-10. doi:10.1016/j.steroids.2014.12.013.

8. Santen RJ, Yue W, Wang J-P. Estrogen metabolites and breast cancer. *Steroids*. 2015;99(Pt A):61-66. doi:10.1016/j.steroids.2014.08.003.

9. Okobia MN, Bunker CH. Estrogen metabolism and breast cancer risk--a review. *Afr J Reprod Health*. 2006;10(1):13-25. http://www.ncbi.nlm.nih.gov/pubmed/16999191. Accessed September 26, 2016.

10. Fuhrman BJ, Moore SC, Matthews CE. Epidemiologic studies of estrogen metabolism and breast cancer. *Steroids*. 2015;99:67-75. doi:10.1016/j.steroids.2015.02.015.

11. Ziegler RG, Fuhrman BJ, Moore SC, Matthews CE. Epidemiologic studies of estrogen metabolism and breast cancer. *Steroids*. 2015;99:67-75. doi:10.1016/j.steroids.2015.02.015.

12. Endogenous Hormones and Breast Cancer Collaborative Group EH and BCC, Key TJ, Appleby PN, et al. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. *Br J Cancer*. 2011;105(5):709-722. doi:10.1038/bjc.2011.254.

13. World Cancer Research Fund / American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective*. Washington DC; 2007.

14. Romaguera D, Vergnaud A-C, Peeters PH, et al. Is concordance with World Cancer Research Fund/American Institute for Cancer Research guidelines for cancer prevention related to subsequent risk of cancer? Results from the EPIC study. *Am J Clin Nutr*. 2012;96(1):150-163. doi:10.3945/ajcn.111.031674.

15. Sánchez-Zamorano LM, Flores-Luna L, Ángeles-Llerenas A, et al. Healthy Lifestyle on the Risk of Breast Cancer. *Cancer Epidemiol Prev Biomarkers*. 2011;20(5).

16. McKenzie F, Ferrari P, Freisling H, et al. Healthy lifestyle and risk of breast cancer among postmenopausal women in the European Prospective Investigation into Cancer and Nutrition cohort study. *Int J cancer*. 2015;136(11):2640-2648. doi:10.1002/ijc.29315.

17. Albuquerque RCR, Baltar VT, Marchioni DML. Breast cancer and dietary patterns: a systematic review. *Nutr Rev*. 2014;72(1):1-17. doi:10.1111/nure.12083.

18. Brennan SF, Cantwell MM, Cardwell CR, Velentzis LS, Woodside J V. Dietary patterns and breast cancer risk: a systematic review and meta-analysis. *Am J Clin Nutr*. 2010;91(5):1294-1302. doi:10.3945/ajcn.2009.28796.

19. Vera-Ramirez L, Ramirez-Tortosa MC, Sanchez-Rovira P, et al. Impact of diet on breast cancer risk: a review of experimental and observational studies. *Crit Rev Food Sci Nutr*. 2013;53(1):49-75. doi:10.1080/10408398.2010.521600.

20. Aune D, Chan DSM, Vieira AR, et al. Fruits, vegetables and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Breast Cancer Res Treat*. 2012;134(2):479-493. doi:10.1007/s10549-012-2118-1.

21. Edefonti V, Randi G, La Vecchia C, Ferraroni M, Decarli A. Dietary patterns and breast cancer: a review with focus on methodological issues. *Nutr Rev*. 2009;67(6):297-314. doi:10.1111/j.1753-4887.2009.00203.x.

22. Thomson CA. Diet and breast cancer: understanding risks and benefits. *Nutr Clin Pract*. 2012;27(5):636-650. doi:10.1177/0884533612454302.

23. Alexander DD, Morimoto LM, Mink PJ, Lowe KA. Summary and meta-analysis of prospective studies of animal fat intake and breast cancer. *Nutr Res Rev*. 2010;23(1):169-179. doi:10.1017/S095442241000003X.

24. *Continuous Update Project Report: Diet, Nutrition, Physical Activity and Breast Cancer.*; 2017. http://www.wcrf.org/breast-cancer-2017. Accessed August 8, 2017.

25. *2015–2020 Dietary Guidelines for Americanss. 8th Edition*. Washington, D.C.; 2015. http://health.gov/dietaryguidelines/2015/guidelines/. Accessed December 15, 2016.

26. Kastorini C-M, Papadakis G, Milionis HJ, et al. Comparative analysis of a-priori and a-posteriori dietary patterns using state-of-the-art classification algorithms: a case/case-control study. *Artif Intell Med*. 2013;59(3):175-183. doi:10.1016/j.artmed.2013.08.005.

27. Harris HR, Bergkvist L, Wolk A. An estrogen-associated dietary pattern and breast cancer risk in the Swedish Mammography Cohort. *Int J cancer*. 2015;137(9):2149-2154. doi:10.1002/ijc.29586.

28. Link LB, Canchola AJ, Bernstein L, et al. Dietary patterns and breast cancer risk in the California Teachers Study cohort. *Am J Clin Nutr*. 2013;98(6):1524-1532. doi:10.3945/ajcn.113.061184.

29. Kulin HE, Bwibo N, Mutie D, Santner SJ. The effect of chronic childhood malnutrition on pubertal growth and development. *Am J Clin Nutr*. 1982;36(3):527-536. http://www.ncbi.nlm.nih.gov/pubmed/7113957. Accessed February 7, 2017.

30. Hunter DJ, Spiegelman D, Adami HO, et al. Non-dietary factors as risk factors for breast cancer, and as effect modifiers of the association of fat intake and risk of breast cancer. *Cancer Causes Control*. 1997;8(1):49-56. http://www.ncbi.nlm.nih.gov/pubmed/9051322. Accessed February 7, 2017.

31. Hoffmann K. Application of a New Statistical Method to Derive Dietary Patterns in Nutritional Epidemiology. *Am J Epidemiol*. 2004;159(10):935-944. doi:10.1093/aje/kwh134.

32. Fung TT, Schulze MB, Hu FB, Hankinson SE, Holmes MD. A dietary pattern derived to correlate with estrogens and risk of postmenopausal breast cancer. *Breast Cancer Res Treat*. 2012;132(3):1157-1162. doi:10.1007/s10549-011-1942-z.

33. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21(6 Suppl):273S-309S. http://www.ncbi.nlm.nih.gov/pubmed/11189684. Accessed October 13, 2016.

34. Sandler D, Hodgson M, Demig-Halverson S, et al. The Sister Study: Baseline methods and participant characteristics. *Environ Health Perspect*. 2017;In press.

35. Fuhrman BJ, Schairer C, Gail MH, et al. Estrogen metabolism and risk of breast

cancer in postmenopausal women. *J Natl Cancer Inst*. 2012;104(4):326-339. doi:10.1093/jnci/djr531.

36. Catsburg C, Kim RS, Kirsh VA, Soskolne CL, Kreiger N, Rohan TE. Dietary patterns and breast cancer risk: a study in 2 cohorts. *Am J Clin Nutr*. 2015;101(4):817-823. doi:10.3945/ajcn.114.097659.

37. Sant M, Allemani C, Sieri S, et al. Salad vegetables dietary pattern protects against HER-2-positive breast cancer: A prospective Italian study. *Int J Cancer*. 2007;121(4):911-914. doi:10.1002/ijc.22714.

38. Sieri S, Krogh V, Pala V, et al. Dietary patterns and risk of breast cancer in the ORDET cohort. *Cancer Epidemiol Biomarkers Prev*. 2004;13(4):567-572. http://www.ncbi.nlm.nih.gov/pubmed/15066921. Accessed November 16, 2016.

39. Thomson CA, Thompson PA. Dietary patterns, risk and prognosis of breast cancer. *Futur Oncol*. 2009;5(8):1257-1269. doi:10.2217/fon.09.86.

40. Farsinejad-Marj M, Talebi S, Ghiyasvand R, Miraghajani M. Adherence to Mediterranean diet and risk of breast cancer in premenopausal and postmenopausal women. *Arch Iran Med*. 2015;18(11):786-792. doi:0151811/AIM.0010.

41. Xu X, Roman JM, Issaq HJ, Keefer LK, Veenstra TD, Ziegler RG. Quantitative measurement of endogenous estrogens and estrogen metabolites in human serum by liquid chromatography-tandem mass spectrometry. *Anal Chem*. 2007;79(20):7813-7821. doi:10.1021/ac070494j.

42. Ryerson AB, Eheman CR, Altekruse SF, et al. Annual Report to the Nation on the Status of Cancer, 1975-2012, featuring the increasing incidence of liver cancer. *Cancer*. 2016;122(9):1312-1337. doi:10.1002/cncr.29936.

43. Eccles SA, Aboagye EO, Ali S, et al. Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. *Breast Cancer Res*. 2013;15(5):R92. doi:10.1186/bcr3493.

44. Mourouti N, Kontogianni MD, Papavagelis C, Panagiotakos DB. Diet and breast cancer: a systematic review. *Int J Food Sci Nutr*. 2015;66(1):1-42. doi:10.3109/09637486.2014.950207.

45. Harnack L, Nicodemus K, Jacobs DR, Folsom AR. An evaluation of the Dietary Guidelines for Americans in relation to cancer occurrence. *Am J Clin Nutr*. 2002;76(4):889-896. http://www.ncbi.nlm.nih.gov/pubmed/12324305. Accessed December 18, 2016.

46. Jankovic N, Geelen A, Winkels RM, et al. Adherence to the WCRF/AICR Dietary Recommendations for Cancer Prevention and Risk of Cancer in Elderly from Europe and the United States: A Meta-Analysis within the CHANCES Project. *Cancer Epidemiol Prev Biomarkers*. 2016.

47. Potentas E, Witkowska AM, Zujko ME. Mediterranean diet for breast cancer prevention and treatment in postmenopausal women. *Menopausal Rev*. 2015;4(4):247-253. doi:10.5114/pm.2015.56381.

48. Toledo E, Salas-Salvadó J, Donat-Vargas C, et al. Mediterranean Diet and Invasive Breast Cancer Risk Among Women at High Cardiovascular Risk in the PREDIMED Trial: A Randomized Clinical Trial. *JAMA Intern Med*. 2015;175(11):1752-1760. doi:10.1001/jamainternmed.2015.4838.

49. Shivappa N, Hébert JR, Rosato V, Montella M, Serraino D, La Vecchia C. Association between the dietary inflammatory index and breast cancer in a large Italian case-control study. *Mol Nutr Food Res*. November 2016. doi:10.1002/mnfr.201600500.

50. Shivappa N, Blair CK, Prizment AE, Jacobs DR, Hébert JR. Prospective study of the dietary inflammatory index and risk of breast cancer in postmenopausal women. *Mol Nutr Food Res*. November 2016. doi:10.1002/mnfr.201600592.

51. Tabung FK, Steck SE, Liese AD, et al. Association between dietary inflammatory potential and breast cancer incidence and death: results from the Women's Health Initiative. *Br J Cancer*. 2016;114(11):1277-1285. doi:10.1038/bjc.2016.98.

52. Purohit V. Moderate alcohol consumption and estrogen levels in postmenopausal women: a review. *Alcohol Clin Exp Res*. 1998;22(5):994-997. http://www.ncbi.nlm.nih.gov/pubmed/9726268. Accessed October 18, 2016.

53. Vicennati V, Garelli S, Rinaldi E, et al. Obesity-related proliferative diseases: the interaction between adipose tissue and estrogens in post-menopausal women. *Horm Mol Biol Clin Investig*. 2015;21(1):75-87. doi:10.1515/hmbci-2015-0002.

54. Kim JH, Cho HT, Kim YJ. The role of estrogen in adipose tissue metabolism: insights into glucose homeostasis regulation. *Endocr J*. 2014;61(11):1055-1067. http://www.ncbi.nlm.nih.gov/pubmed/25109846. Accessed October 24, 2016.

55. Gruber CJ, Tschugguel W, Schneeberger C, Huber JC. Production and actions of estrogens. *N Engl J Med*. 2002;346(5):340-352. doi:10.1056/NEJMra000471.

56. Nelson LR, Bulun SE. Estrogen production and action. *J Am Acad Dermatol*. 2001;45(3 Suppl):S116-24. http://www.ncbi.nlm.nih.gov/pubmed/11511861. Accessed October 24, 2016.

57. Key T, Appleby P, Barnes I, Reeves G, Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst*. 2002;94(8):606-616. http://www.ncbi.nlm.nih.gov/pubmed/11959894. Accessed August 24, 2016.

58. Folkerd E, Dowsett M. Sex hormones and breast cancer risk and prognosis. *Breast*.

2013;22 Suppl 2:S38-43. doi:10.1016/j.breast.2013.07.007.

59. Woolcott CG, Shvetsov YB, Stanczyk FZ, et al. Plasma sex hormone concentrations and breast cancer risk in an ethnically diverse population of postmenopausal women: the Multiethnic Cohort Study. *Endocr Relat Cancer*. 2010;17(1):125-134. doi:10.1677/ERC-09-0211.

60. Stanczyk FZ, Mathews BW, Sherman ME. Relationships of sex steroid hormone levels in benign and cancerous breast tissue and blood: A critical appraisal of current science. *Steroids*. 2015;99(Pt A):91-102. doi:10.1016/j.steroids.2014.12.011.

61. Dorgan JF, Stanczyk FZ, Longcope C, et al. Relationship of serum dehydroepiandrosterone (DHEA), DHEA sulfate, and 5-androstene-3 beta, 17 beta-diol to risk of breast cancer in postmenopausal women. *Cancer Epidemiol Biomarkers Prev*. 1997;6(3):177-181. http://www.ncbi.nlm.nih.gov/pubmed/9138660. Accessed August 27, 2016.

62. Schmidt S, Monk JM, Robinson LE, Mourtzakis M. The integrative role of leptin, oestrogen and the insulin family in obesity-associated breast cancer: potential effects of exercise. *Obes Rev*. 2015;16(6):473-487. doi:10.1111/obr.12281.

63. Lorincz AM, Sukumar S. Molecular links between obesity and breast cancer. *Endocr Relat Cancer*. 2006;13(2):279-292. doi:10.1677/erc.1.00729.

64. Jardé T, Perrier S, Vasson M-P, Caldefie-Chézet F. Molecular mechanisms of leptin and adiponectin in breast cancer. *Eur J Cancer*. 2011;47(1):33-43. doi:10.1016/j.ejca.2010.09.005.

65. Tworoger SS, Hankinson SE. Prolactin and breast cancer etiology: an epidemiologic perspective. *J Mammary Gland Biol Neoplasia*. 2008;13(1):41-53. doi:10.1007/s10911-008-9063-y.

66. Carey LA, Perou CM, Livasy CA, et al. Race, Breast Cancer Subtypes, and Survival in the Carolina Breast Cancer Study. *JAMA*. 2006;295(21):2492. doi:10.1001/jama.295.21.2492.

67. Hulka BS, Moorman PG. Breast cancer: hormones and other risk factors. *Maturitas*. 2001;38(1):103-13-6. http://www.ncbi.nlm.nih.gov/pubmed/11311599. Accessed October 26, 2016.

68. Ziegler RG, Rossi SC, Fears TR, et al. Quantifying estrogen metabolism: an evaluation of the reproducibility and validity of enzyme immunoassays for 2-hydroxyestrone and 16alpha-hydroxyestrone in urine. *Environ Health Perspect*. April 1997:607-614. http://www.ncbi.nlm.nih.gov/pubmed/9168003. Accessed December 8, 2016.

69. Key TJ, Appleby PN, Reeves GK, et al. Steroid hormone measurements from

different types of assays in relation to body mass index and breast cancer risk in postmenopausal women: Reanalysis of eighteen prospective studies. *Steroids*. 2015;99(Pt A):49-55. doi:10.1016/j.steroids.2014.09.001.

70. Xu X, Veenstra TD, Fox SD, et al. Measuring fifteen endogenous estrogens simultaneously in human urine by high-performance liquid chromatography-mass spectrometry. *Anal Chem*. 2005;77(20):6646-6654. doi:10.1021/ac050697c.

71. Spierto FW, Gardner F, Smith SJ. Evaluation of an EIA method for measuring serum levels of the estrogen metabolite 2-hydroxyestrone in adults. *Steroids*. 2001;66(1):59-62. http://www.ncbi.nlm.nih.gov/pubmed/11090660. Accessed December 8, 2016.

72. Faupel-Badger JM, Fuhrman BJ, Xu X, et al. Comparison of liquid chromatography-tandem mass spectrometry, RIA, and ELISA methods for measurement of urinary estrogens. *Cancer Epidemiol Biomarkers Prev*. 2010;19(1):292-300. doi:10.1158/1055-9965.EPI-09-0643.

73. Fuhrman BJ, Xu X, Falk RT, et al. Assay Reproducibility and Interindividual Variation for 15 Serum Estrogens and Estrogen Metabolites Measured by Liquid Chromatography-Tandem Mass Spectrometry. *Cancer Epidemiol Biomarkers Prev*. 2014;23(12):2649-2657. doi:10.1158/1055-9965.EPI-14-0438.

74. Falk RT, Xu X, Keefer L, Veenstra TD, Ziegler RG. A Liquid Chromatography-Mass Spectrometry Method for the Simultaneous Measurement of 15 Urinary Estrogens and Estrogen Metabolites: Assay Reproducibility and Interindividual Variability. *Cancer Epidemiol Biomarkers Prev*. 2008;17(12):3411-3418. doi:10.1158/1055-9965.EPI-08-0355.

75. Lee JS, Ettinger B, Stanczyk FZ, et al. Comparison of methods to measure low serum estradiol levels in postmenopausal women. *J Clin Endocrinol Metab*. 2006;91(10):3791-3797. doi:10.1210/jc.2005-2378.

76. Rosner W, Hankinson SE, Sluss PM, Vesper HW, Wierman ME. Challenges to the Measurement of Estradiol: An Endocrine Society Position Statement. *J Clin Endocrinol Metab*. 2013;98(4):1376-1387. doi:10.1210/jc.2012-3780.

77. Williams RH, Melmed S. *Williams Textbook of Endocrinology*. Elsevier/Saunders; 2011.

78. RYAN KJ. Biological aromatization of steroids. *J Biol Chem*. 1959;234(2):268-272. http://www.ncbi.nlm.nih.gov/pubmed/13630892. Accessed February 7, 2017.

79. Dallal CM, Tice JA, Buist DSM, et al. Estrogen metabolism and breast cancer risk among postmenopausal women: a case-cohort study within B~FIT. *Carcinogenesis*. 2014;35(2):346-355. doi:10.1093/carcin/bgt367.

80. Falk RT, Brinton LA, Dorgan JF, et al. Relationship of serum estrogens and

estrogen metabolites to postmenopausal breast cancer risk: a nested case-control study. *Breast Cancer Res*. 2013;15(2):R34. doi:10.1186/bcr3416.

81. Yager JD, Davidson NE. Estrogen Carcinogenesis in Breast Cancer. *N Engl J Med*. 2006;354(3):270-282. doi:10.1056/NEJMra050776.

82. Yager JD, Davidson NE. Estrogen carcinogenesis in breast cancer. *N Engl J Med*. 2006;354(3):270-282. doi:10.1056/NEJMra050776.

83. Pike MC, Spicer D V, Dahmoush L, Press MF. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiol Rev*. 1993;15(1):17-35. http://www.ncbi.nlm.nih.gov/pubmed/8405201. Accessed August 31, 2016.

84. Telang NT, Suto A, Wong GY, Osborne MP, Bradlow HL. Induction by estrogen metabolite 16 alpha-hydroxyestrone of genotoxic damage and aberrant proliferation in mouse mammary epithelial cells. *J Natl Cancer Inst*. 1992;84(8):634-638. http://www.ncbi.nlm.nih.gov/pubmed/1556774. Accessed December 15, 2016.

85. Arnold M, Razum O, Coebergh J-W. Cancer risk diversity in non-western migrants to Europe: An overview of the literature. *Eur J Cancer*. 2010;46(14):2647-2659. doi:10.1016/j.ejca.2010.07.050.

86. Parkin DM, Khlat M. Studies of cancer in migrants: rationale and methodology. *Eur J Cancer*. 1996;32A(5):761-771. http://www.ncbi.nlm.nih.gov/pubmed/9081351. Accessed October 14, 2016.

87. Ziegler RG, Hoover RN, Pike MC, et al. Migration patterns and breast cancer risk in Asian-American women. *J Natl Cancer Inst*. 1993;85(22):1819-1827. http://www.ncbi.nlm.nih.gov/pubmed/8230262. Accessed October 14, 2016.

88. Goldin BR, Adlercreutz H, Gorbach SL, et al. The relationship between estrogen levels and diets of Caucasian American and Oriental immigrant women. *Am J Clin Nutr*. 1986;44(6):945-953. http://www.ncbi.nlm.nih.gov/pubmed/3024478. Accessed December 13, 2016.

89. Goldin BR, Adlercreutz H, Gorbach SL, et al. Estrogen Excretion Patterns and Plasma Levels in Vegetarian and Omnivorous Women. *N Engl J Med*. 1982;307(25):1542-1547. doi:10.1056/NEJM198212163072502.

90. Brinkman MT, Baglietto L, Krishnan K, et al. Consumption of animal products, their nutrient components and postmenopausal circulating steroid hormone concentrations. *Eur J Clin Nutr*. 2010;64(2):176-183. doi:10.1038/ejcn.2009.129.

91. Furnari C, Maroun D, Gyawali S, Snyder BW, Davis AM. Lack of biologically active estrogens in commercial cow milk. *J Dairy Sci*. 2012;95(1):9-14. doi:10.3168/jds.2011-4365.

92. Consumption of Cow's Milk and Possible Risk of Breast Cancer. *Breast Care (Basel)*. 2010;5(1):44-46. doi:10.1159/000277938.

93. Gann PH, Chatterton RT, Gapstur SM, et al. The effects of a low-fat/high-fiber diet on sex hormone levels and menstrual cycling in premenopausal women. *Cancer*. 2003;98(9):1870-1879. doi:10.1002/cncr.11735.

94. Aubertin-Leheudre M, Gorbach S, Woods M, Dwyer JT, Goldin B, Adlercreutz H. Fat/fiber intakes and sex hormones in healthy premenopausal women in USA. *J Steroid Biochem Mol Biol*. 2008;112(1-3):32-39. doi:10.1016/j.jsbmb.2008.08.002.

95. Goldin BR, Woods MN, Spiegelman DL, et al. The effect of dietary fat and fiber on serum estrogen concentrations in premenopausal women under controlled dietary conditions. *Cancer*. 1994;74(3 Suppl):1125-1131. http://www.ncbi.nlm.nih.gov/pubmed/8039147. Accessed December 14, 2016.

96. Rose DP, Goldman M, Connolly JM, Strong LE. High-fiber diet reduces serum estrogen concentrations in premenopausal women. *Am J Clin Nutr*. 1991;54(3):520-525. http://www.ncbi.nlm.nih.gov/pubmed/1652197. Accessed December 14, 2016.

97. Woods MN, Barnett JB, Spiegelman D, et al. Hormone levels during dietary changes in premenopausal African-American women. *J Natl Cancer Inst*. 1996;88(19):1369-1374. http://www.ncbi.nlm.nih.gov/pubmed/8827014. Accessed December 14, 2016.

98. Oh H, Smith-Warner SA, Tamimi RM, et al. Dietary Fat and Fiber Intakes Are Not Associated with Patterns of Urinary Estrogen Metabolites in Premenopausal Women. *J Nutr*. 2015;145(9):2109-2116. doi:10.3945/jn.115.212779.

99. Rock CL, Flatt SW, Thomson CA, et al. Effects of a high-fiber, low-fat diet intervention on serum concentrations of reproductive steroid hormones in women with a history of breast cancer. *J Clin Oncol*. 2004;22(12):2379-2387. doi:10.1200/JCO.2004.09.025.

100. Monroe KR, Murphy SP, Henderson BE, et al. Dietary fiber intake and endogenous serum hormone levels in naturally postmenopausal Mexican American women: the Multiethnic Cohort Study. *Nutr Cancer*. 2007;58(2):127-135. doi:10.1080/01635580701327935.

101. Shultz TD, Howie BJ. In vitro binding of steroid hormones by natural and purified fibers. *Nutr Cancer*. 1986;8(2):141-147. doi:10.1080/01635588609513887.

102. Holmes MD, Spiegelman D, Willett WC, et al. Dietary fat intake and endogenous sex steroid hormone levels in postmenopausal women. *J Clin Oncol*. 2000;18(21):3668-3676. doi:10.1200/jco.2000.18.21.3668.

103. Katsouyanni K, Boyle P, Trichopoulos D. Diet and urine estrogens among postmenopausal women. *Oncology*. 1991;48(6):490-494. http://www.ncbi.nlm.nih.gov/pubmed/1749588. Accessed December 14, 2016.

104. Newcomb PA, Klein R, Klein BE, et al. Association of dietary and life-style factors with sex hormones in postmenopausal women. *Epidemiology*. 1995;6(3):318-321. http://www.ncbi.nlm.nih.gov/pubmed/7619943. Accessed December 14, 2016.

105. Nagata C, Nagao Y, Shibuya C, Kashiki Y, Shimizu H. Fat intake is associated with serum estrogen and androgen concentrations in postmenopausal Japanese women. *J Nutr*. 2005;135(12):2862-2865. http://www.ncbi.nlm.nih.gov/pubmed/16317133. Accessed December 14, 2016.

106. Aubertin-Leheudre M, Hamalainen E, Adlercreutz H. Diets and Hormonal Levels in Postmenopausal Women With or Without Breast Cancer. *Nutr Cancer*. 2011;63(4):514-524. doi:10.1080/01635581.2011.538487.

107. Prentice R, Thompson D, Clifford C, Gorbach S, Goldin B, Byar D. Dietary fat reduction and plasma estradiol concentration in healthy postmenopausal women. The Women's Health Trial Study Group. *J Natl Cancer Inst*. 1990;82(2):129-134. http://www.ncbi.nlm.nih.gov/pubmed/2294222. Accessed December 13, 2016.

108. Heber D, Ashley JM, Leaf DA, Barnard RJ. Reduction of serum estradiol in postmenopausal women given free access to low-fat high-carbohydrate diet. *Nutrition*. 7(2):137-9-40. http://www.ncbi.nlm.nih.gov/pubmed/1666318. Accessed December 13, 2016.

109. Young LR, Raatz SK, Thomas W, Redmon J, Kurzer MS. Total dietary fat and omega-3 fatty acids have modest effects on urinary sex hormones in postmenopausal women. *Nutr Metab (Lond)*. 2013;10(1):36. doi:10.1186/1743-7075-10-36.

110. Prentice RL, Caan B, Chlebowski RT, et al. Low-fat dietary pattern and risk of invasive breast cancer: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA*. 2006;295(6):629-642. doi:10.1001/jama.295.6.629.

111. Wu AH, Pike MC, Stram DO. Meta-analysis: dietary fat intake, serum estrogen levels, and the risk of breast cancer. *J Natl Cancer Inst*. 1999;91(6):529-534. doi:10.1093/JNCI/91.6.529.

112. Wu AH, Stanczyk FZ, Seow A, Lee H-P, Yu MC. Soy Intake and Other Lifestyle Determinants of Serum Estrogen Levels among Postmenopausal Chinese Women in Singapore. *Cancer Epidemiol Prev Biomarkers*. 2002;11(9).

113. London S, Willett W, Longcope C, McKinlay S. Alcohol and other dietary factors in relation to serum hormone concentrations in women at climacteric. *Am J Clin*

Nutr. 1991;53(1):166-171. http://www.ncbi.nlm.nih.gov/pubmed/1845789. Accessed December 14, 2016.

114. Yoshimura N, Kasamatsu T, Sakata K, Hashimoto T, Cooper C. The relationship between endogenous estrogen, sex hormone-binding globulin, and bone loss in female residents of a rural Japanese community: the Taiji Study. *J Bone Miner Metab*. 2002;20(5):303-310. doi:10.1007/s007740200044.

115. Michnovicz JJ, Bradlow HL. Induction of Estradiol Metabolism by Dietary Indole-3-carbinol in Humans. *JNCI J Natl Cancer Inst*. 1990;82(11):947-949. doi:10.1093/jnci/82.11.947.

116. BRADLOW HL, SEPKOVIC DW, TELANG NT, OSBORNE MP. Indole-3-carbinol. *Ann N Y Acad Sci*. 1995;768(1):180-200. doi:10.1111/j.1749-6632.1995.tb12121.x.

117. Michnovicz JJ, Adlercreutz H, Bradlow HL. Changes in Levels of Urinary Estrogen Metabolites After Oral Indole-3-Carbinol Treatment in Humans. *JNCI J Natl Cancer Inst*. 1997;89(10):718-723. doi:10.1093/jnci/89.10.718.

118. Fowke JH, Longcope C, Hebert JR. Brassica Vegetable Consumption Shifts Estrogen Metabolism in Healthy Postmenopausal Women. *Cancer Epidemiol Prev Biomarkers*. 2000;9(8).

119. Morrison J, Mutell D, Pollock TA, Redmond E, Bralley JA, Lord RS. Effects of dried cruciferous powder on raising 2/16 hydroxyestrogen ratios in premenopausal women. *Altern Ther Health Med*. 15(2):52-53. http://www.ncbi.nlm.nih.gov/pubmed/19284183. Accessed December 14, 2016.

120. Sowers MR, Crawford S, McConnell DS, et al. Selected diet and lifestyle factors are associated with estrogen metabolites in a multiracial/ethnic population of women. *J Nutr*. 2006;136(6):1588-1595. http://www.ncbi.nlm.nih.gov/pubmed/16702326. Accessed December 14, 2016.

121. Latté KP, Appel K-E, Lampen A. Health benefits and possible risks of broccoli - an overview. *Food Chem Toxicol*. 2011;49(12):3287-3309. doi:10.1016/j.fct.2011.08.019.

122. Yoshida M, Katashima S, Ando J, et al. Dietary indole-3-carbinol promotes endometrial adenocarcinoma development in rats initiated with N-ethyl-N'-nitro-N-nitrosoguanidine, with induction of cytochrome P450s in the liver and consequent modulation of estrogen metabolism. *Carcinogenesis*. 2004;25(11):2257-2264. doi:10.1093/carcin/bgh225.

123. Travis RC, Allen NE, Appleby PN, Spencer EA, Roddam AW, Key TJ. A prospective study of vegetarianism and isoflavone intake in relation to breast cancer risk in British women. *Int J cancer*. 2008;122(3):705-710. doi:10.1002/ijc.23141.

124. Adlercreutz H, Mazur W. Phyto-oestrogens and Western diseases. *Ann Med*. 1997;29(2):95-120. http://www.ncbi.nlm.nih.gov/pubmed/9187225. Accessed December 15, 2016.

125. Wu AH, Yu MC, Tseng C-C, Pike MC. Epidemiology of soy exposures and breast cancer risk. *Br J Cancer*. 2008;98(1):9-14. doi:10.1038/sj.bjc.6604145.

126. Xu X, Duncan AM, Merz BE, Kurzer MS. Effects of soy isoflavones on estrogen and phytoestrogen metabolism in premenopausal women. *Cancer Epidemiol Biomarkers Prev*. 1998;7(12):1101-1108. http://www.ncbi.nlm.nih.gov/pubmed/9865428. Accessed December 15, 2016.

127. Morimoto Y, Conroy SM, Pagano IS, et al. Urinary Estrogen Metabolites During a Randomized Soy Trial. *Nutr Cancer*. 2012;64(2):307-314. doi:10.1080/01635581.2012.648819.

128. Baird DD, Umbach DM, Lansdell L, et al. Dietary intervention study to assess estrogenicity of dietary soy among postmenopausal women. *J Clin Endocrinol Metab*. 1995;80(5):1685-1690. doi:10.1210/jcem.80.5.7745019.

129. Hooper L, Ryder JJ, Kurzer MS, et al. Effects of soy protein and isoflavones on circulating hormone concentrations in pre- and post-menopausal women: a systematic review and meta-analysis. *Hum Reprod Update*. 2009;15(4):423-440. doi:10.1093/humupd/dmp010.

130. Fuhrman BJ, Pfeiffer R, Xu X, et al. Soy Intake is Associated with Increased 2-Hydroxylation and Decreased 16 -Hydroxylation of Estrogens in Asian-American Women. *Cancer Epidemiol Biomarkers Prev*. 2009;18(10):2751-2760. doi:10.1158/1055-9965.EPI-09-0388.

131. Verkasalo PK, Appleby PN, Davey GK, Key TJ. Soy milk intake and plasma sex hormones: a cross-sectional study in pre- and postmenopausal women (EPIC-Oxford). *Nutr Cancer*. 2001;40(2):79-86. doi:10.1207/S15327914NC402_1.

132. Low Y-L, Dunning AM, Dowsett M, et al. Phytoestrogen Exposure Is Associated with Circulating Sex Hormone Levels in Postmenopausal Women and Interact with ESR1 and NR1I2 Gene Variants. *Cancer Epidemiol Biomarkers Prev*. 2007;16(5):1009-1016. doi:10.1158/1055-9965.EPI-06-0899.

133. Liu Y, Nguyen N, Colditz GA. Links between alcohol consumption and breast cancer: a look at the evidence. *Womens Health (Lond Engl)*. 2015;11(1):65-77. doi:10.2217/whe.14.62.

134. Wu AH, Vigen C, Razavi P, Tseng C-C, Stanczyk FZ. Alcohol and Breast Cancer Risk among Asian American Women in Los Angeles County. *Breast Cancer Res*. 2012;14(6):R151. doi:10.1186/bcr3363.

135. Singletary KW, Gapstur SM. Alcohol and breast cancer: review of epidemiologic

and experimental evidence and potential mechanisms. *JAMA*. 2001;286(17):2143-2151. http://www.ncbi.nlm.nih.gov/pubmed/11694156. Accessed October 18, 2016.

136. Rinaldi S, Peeters PHM, Bezemer ID, et al. Relationship of alcohol intake and sex steroid concentrations in blood in pre- and post-menopausal women: the European Prospective Investigation into Cancer and Nutrition. *Cancer Causes Control*. 2006;17(8):1033-1043. doi:10.1007/s10552-006-0041-7.

137. Reichman ME, Judd JT, Longcope C, et al. Effects of alcohol consumption on plasma and urinary hormone concentrations in premenopausal women. *J Natl Cancer Inst*. 1993;85(9):722-727. http://www.ncbi.nlm.nih.gov/pubmed/8478958. Accessed December 16, 2016.

138. Muti P, Trevisan M, Micheli A, et al. Alcohol consumption and total estradiol in premenopausal women. *Cancer Epidemiol Biomarkers Prev*. 1998;7(3):189-193. http://www.ncbi.nlm.nih.gov/pubmed/9521430. Accessed December 16, 2016.

139. Sarkola T, Makisalo H, Fukunaga T, Eriksson CJP. Acute Effect of Alcohol on Estradiol, Estrone, Progesterone, Prolactin, Cortisol, and Luteinizing Hormone in Premenopausal Women. *Alcohol Clin Exp Res*. 1999;23(6):976-982. doi:10.1111/j.1530-0277.1999.tb04215.x.

140. Hankinson SE, Willett WC, Manson JE, et al. Alcohol, height, and adiposity in relation to estrogen and prolactin levels in postmenopausal women. *J Natl Cancer Inst*. 1995;87(17):1297-1302. doi:10.1093/JNCI/87.17.1297.

141. Onland-Moret NC, Peeters PHM, van der Schouw YT, Grobbee DE, van Gils CH. Alcohol and Endogenous Sex Steroid Levels in Postmenopausal Women: A Cross-Sectional Study. *J Clin Endocrinol Metab*. 2005;90(3):1414-1419. doi:10.1210/jc.2004-0614.

142. Gavaler JS, Love K, Van Thiel D, et al. An international study of the relationship between alcohol consumption and postmenopausal estradiol levels. *Alcohol Alcohol Suppl*. 1991;1:327-330. http://www.ncbi.nlm.nih.gov/pubmed/1845556. Accessed December 16, 2016.

143. Sierksma A, Sarkola T, Eriksson CJP, van der Gaag MS, Grobbee DE, Hendriks HFJ. Effect of moderate alcohol consumption on plasma dehydroepiandrosterone sulfate, testosterone, and estradiol levels in middle-aged men and postmenopausal women: a diet-controlled intervention study. *Alcohol Clin Exp Res*. 2004;28(5):780-785. http://www.ncbi.nlm.nih.gov/pubmed/15166654. Accessed December 16, 2016.

144. Fung TT, Hu FB, Barbieri RL, Willett WC, Hankinson SE. Dietary patterns, the Alternate Healthy Eating Index and plasma sex hormone concentrations in postmenopausal women. *Int J Cancer*. 2007;121(4):803-809. doi:10.1002/ijc.22728.

145. Sánchez-Zamorano LM, Flores-Luna L, Angeles-Llerenas A, et al. The Western dietary pattern is associated with increased serum concentrations of free estradiol in postmenopausal women: implications for breast cancer prevention. *Nutr Res*. 2016;36(8):845-854. doi:10.1016/j.nutres.2016.04.008.

146. Hirko KA, Spiegelman D, Barnett JB, et al. Dietary Patterns and Plasma Sex Hormones, Prolactin, and Sex Hormone-Binding Globulin in Premenopausal Women. *Cancer Epidemiol Biomarkers Prev*. 2016;25(5):791-798. doi:10.1158/1055-9965.EPI-15-1019.

147. Carruba G, Granata OM, Pala V, et al. A traditional Mediterranean diet decreases endogenous estrogens in healthy postmenopausal women. *Nutr Cancer*. 2006;56(2):253-259. doi:10.1207/s15327914nc5602_18.

148. Anand P, Kunnumakkara AB, Kunnumakara AB, et al. Cancer is a preventable disease that requires major lifestyle changes. *Pharm Res*. 2008;25(9):2097-2116. doi:10.1007/s11095-008-9661-9.

149. Willett WC. Diet, nutrition, and avoidable cancer. *Environ Health Perspect*. November 1995:165-170. http://www.ncbi.nlm.nih.gov/pubmed/8741778. Accessed October 14, 2016.

150. Aune D, Chan DSM, Greenwood DC, et al. Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Ann Oncol*. 2012;23(6):1394-1402. doi:10.1093/annonc/mdr589.

151. Emaus MJ, Peeters PH, Bakker MF, et al. Vegetable and fruit consumption and the risk of hormone receptor-defined breast cancer in the EPIC cohort. *Am J Clin Nutr*. 2016;103(1):168-177. doi:10.3945/ajcn.114.101436.

152. Masala G, Assedi M, Bendinelli B, et al. Fruit and vegetables consumption and breast cancer risk: the EPIC Italy study. *Breast Cancer Res Treat*. 2012;132(3):1127-1136. doi:10.1007/s10549-011-1939-7.

153. Jung S, Spiegelman D, Baglietto L, et al. Fruit and Vegetable Intake and Risk of Breast Cancer by Hormone Receptor Status. *JNCI J Natl Cancer Inst*. 2013;105(3):219-236. doi:10.1093/jnci/djs635.

154. Link LB, Potter JD. Raw versus cooked vegetables and cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2004;13(9):1422-1435. http://www.ncbi.nlm.nih.gov/pubmed/15342442. Accessed February 8, 2017.

155. Cui Y, Rohan TE. Vitamin D, Calcium, and Breast Cancer Risk: A Review. *Cancer Epidemiol Biomarkers Prev*. 2006;15(8):1427-1437. doi:10.1158/1055-9965.EPI-06-0075.

156. Fay MP, Freedman LS, Clifford CK, Midthune DN. Effect of different types and amounts of fat on the development of mammary tumors in rodents: a review.

Cancer Res. 1997;57(18):3979-3988. http://www.ncbi.nlm.nih.gov/pubmed/9307282. Accessed October 19, 2016.

157. Freedman LS, Clifford C, Messina M. Analysis of dietary fat, calories, body weight, and the development of mammary tumors in rats and mice: a review. *Cancer Res*. 1990;50(18):5710-5719. http://www.ncbi.nlm.nih.gov/pubmed/2203521. Accessed October 19, 2016.

158. Vera-Ramirez L, Ramirez-Tortosa MC, Sanchez-Rovira P, et al. Impact of diet on breast cancer risk: a review of experimental and observational studies. *Crit Rev Food Sci Nutr*. 2013;53(1):49-75. doi:10.1080/10408398.2010.521600.

159. *Continuous Update Project. Food, Nutrition, Physical Activity, and the Prevention of Breast Cancer*.; 2010.

160. Löf M, Sandin S, Lagiou P, et al. Dietary fat and breast cancer risk in the Swedish women's lifestyle and health cohort. *Br J Cancer*. 2007;97(11):1570-1576. doi:10.1038/sj.bjc.6604033.

161. Abel S, Riedel S, Gelderblom WCA. Dietary PUFA and cancer. *Proc Nutr Soc*. 2014;73(3):361-367. doi:10.1017/S0029665114000585.

162. Khodarahmi M, Azadbakht L. The association between different kinds of fat intake and breast cancer risk in women. *Int J Prev Med*. 2014;5(1):6-15. http://www.ncbi.nlm.nih.gov/pubmed/24554986. Accessed December 16, 2016.

163. Sieri S, Krogh V, Ferrari P, et al. Dietary fat and breast cancer risk in the European Prospective Investigation into Cancer and Nutrition. *Am J Clin Nutr*. 2008;88(5):1304-1312. http://www.ncbi.nlm.nih.gov/pubmed/18996867. Accessed December 16, 2016.

164. Saadatian-Elahi M, Norat T, Goudable J, Riboli E. Biomarkers of dietary fatty acid intake and the risk of breast cancer: A meta-analysis. *Int J Cancer*. 2004;111(4):584-591. doi:10.1002/ijc.20284.

165. Voorrips LE, Brants HAM, Kardinaal AFM, Hiddink GJ, van den Brandt PA, Goldbohm RA. Intake of conjugated linoleic acid, fat, and other fatty acids in relation to postmenopausal breast cancer: the Netherlands Cohort Study on Diet and Cancer. *Am J Clin Nutr*. 2002;76(4):873-882. http://www.ncbi.nlm.nih.gov/pubmed/12324303. Accessed December 16, 2016.

166. Wolk A, Bergström R, Hunter D, et al. A prospective study of association of monounsaturated fat and other types of fat with risk of breast cancer. *Arch Intern Med*. 1998;158(1):41-45. http://www.ncbi.nlm.nih.gov/pubmed/9437377. Accessed December 16, 2016.

167. Mourouti N, Panagiotakos DB. Soy food consumption and breast cancer. *Maturitas*. 2013;76(2):118-122. doi:10.1016/j.maturitas.2013.07.006.

168. Chen M, Rao Y, Zheng Y, et al. Association between Soy Isoflavone Intake and Breast Cancer Risk for Pre- and Post-Menopausal Women: A Meta-Analysis of Epidemiological Studies. Ahmad A, ed. *PLoS One*. 2014;9(2):e89288. doi:10.1371/journal.pone.0089288.

169. Dong J-Y, Qin L-Q. Soy isoflavones consumption and risk of breast cancer incidence or recurrence: a meta-analysis of prospective studies. *Breast Cancer Res Treat*. 2011;125(2):315-323. doi:10.1007/s10549-010-1270-8.

170. Nagata C, Mizoue T, Tanaka K, et al. Soy Intake and Breast Cancer Risk: An Evaluation Based on a Systematic Review of Epidemiologic Evidence Among the Japanese Population. *Jpn J Clin Oncol*. 2014;44(3):282-295. doi:10.1093/jjco/hyt203.

171. Dietary Guidelines Advisory Committee. *Scientific Report of the 2015 Dietary Guidelines Advisory Committee*. Washington D.C.; 2015.

172. Mullie P, Koechlin A, Boniol M, Autier P, Boyle P. Relation between Breast Cancer and High Glycemic Index or Glycemic Load: A Meta-analysis of Prospective Cohort Studies. *Crit Rev Food Sci Nutr*. 2016;56(1):152-159. doi:10.1080/10408398.2012.718723.

173. Willett W. *Nutritional Epidemiology*. Oxford University Press; 2012. doi:10.1093/acprof:oso/9780199754038.001.0001.

174. Krebs-Smith SM, Subar AF, Reedy J. Examining Dietary Patterns in Relation to Chronic Disease. *Circulation*. 2015;132(9).

175. Shin S, Saito E, Inoue M, et al. Dietary pattern and breast cancer risk in Japanese women: the Japan Public Health Center-based Prospective Study (JPHC Study). *Br J Nutr*. 2016;115(10):1769-1779. doi:10.1017/S0007114516000684.

176. Karimi Z, Jessri M, Houshiar-Rad A, Mirzaei H-R, Rashidkhani B. Dietary patterns and breast cancer risk among women. *Public Health Nutr*. 2014;17(5):1098-1106. doi:10.1017/S1368980013001018.

177. Cottet V, Touvier M, Fournier A, et al. Postmenopausal Breast Cancer Risk and Dietary Patterns in the E3N-EPIC Prospective Cohort Study. *Am J Epidemiol*. 2009;170(10):1257-1267. doi:10.1093/aje/kwp257.

178. Castelló A, Boldo E, Pérez-Gómez B, et al. Adherence to the Western, Prudent and Mediterranean dietary patterns and breast cancer risk: MCC-Spain study. *Maturitas*. 2017;103:8-15. doi:10.1016/j.maturitas.2017.06.020.

179. Cui X, Dai Q, Tseng M, Shu X-O, Gao Y-T, Zheng W. Dietary Patterns and Breast Cancer Risk in the Shanghai Breast Cancer Study. *Cancer Epidemiol Biomarkers & Prev*. 2007;16(7):1443-1448. doi:10.1158/1055-9965.EPI-07-0059.

180. Adebamowo CA, Hu FB, Cho E, Spiegelman D, Holmes MD, Willett WC. Dietary Patterns and the Risk of Breast Cancer. *Ann Epidemiol*. 2005;15(10):789-795. doi:10.1016/j.annepidem.2005.01.008.

181. Baglietto L, Krishnan K, Severi G, et al. Dietary patterns and risk of breast cancer. *Br J Cancer*. 2011;104(3):524-531. doi:10.1038/sj.bjc.6606044.

182. Cade JE, Taylor EF, Burley VJ, Greenwood DC. Common Dietary Patterns and Risk of Breast Cancer: Analysis From the United Kingdom Women's Cohort Study. *Nutr Cancer*. 2010;62(3):300-306. doi:10.1080/01635580903441246.

183. Agurs-Collins T, Rosenberg L, Makambi K, Palmer JR, Adams-Campbell L. Dietary patterns and breast cancer risk in women participating in the Black Women's Health Study. *Am J Clin Nutr*. 2009;90(3):621-628. doi:10.3945/ajcn.2009.27666.

184. Mills PK, Beeson WL, Phillips RL, Fraser GE. Dietary habits and breast cancer incidence among Seventh-day Adventists. *Cancer*. 1989;64(3):582-590. http://www.ncbi.nlm.nih.gov/pubmed/2743252. Accessed November 16, 2016.

185. Schulz M, Hoffmann K, Weikert C, Nöthlings U, Schulze MB, Boeing H. Identification of a dietary pattern characterized by high-fat food choices associated with increased risk of breast cancer: the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. *Br J Nutr*. 2008;100(5):942-946. doi:10.1017/S0007114508966149.

186. Hirko KA, Willett WC, Hankinson SE, et al. Healthy dietary patterns and risk of breast cancer by molecular subtype. *Breast Cancer Res Treat*. 2016;155(3):579-588. doi:10.1007/s10549-016-3706-2.

187. Nkondjock A, Ghadirian P. Diet quality and BRCA-associated breast cancer risk. *Breast Cancer Res Treat*. 2007;103(3):361-369. doi:10.1007/s10549-006-9371-0.

188. Fung TT, Hu FB, McCullough ML, Newby PK, Willett WC, Holmes MD. Diet quality is associated with the risk of estrogen receptor-negative breast cancer in postmenopausal women. *J Nutr*. 2006;136(2):466-472. http://www.ncbi.nlm.nih.gov/pubmed/16424129. Accessed September 28, 2016.

189. Panico S, Mattiello A, Panico C, Chiodini P. Mediterranean dietary pattern and chronic diseases. *Cancer Treat Res*. 2014;159:69-81. doi:10.1007/978-3-642-38007-5_5.

190. Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. *Cancer Med*. 2015;4(12):1933-1947. doi:10.1002/cam4.539.

191. Appel LJ, Moore TJ, Obarzanek E, et al. A Clinical Trial of the Effects of Dietary Patterns on Blood Pressure. *N Engl J Med*. 1997;336(16):1117-1124.

doi:10.1056/NEJM199704173361601.

192. Ge I, Rudolph A, Shivappa N, Flesch-Janys D, Hébert JR, Chang-Claude J. Dietary inflammation potential and postmenopausal breast cancer risk in a German case-control study. *The Breast*. 2015;24(4):491-496. doi:10.1016/j.breast.2015.04.012.

193. Shivappa N, Sandin S, Löf M, Hébert JR, Adami H-O, Weiderpass E. Prospective study of dietary inflammatory index and risk of breast cancer in Swedish women. *Br J Cancer*. 2015;113(7):1099-1103. doi:10.1038/bjc.2015.304.

194. Surakasula A, Nagarjunapu GC, Raghavaiah K V. A comparative study of pre- and post-menopausal breast cancer: Risk factors, presentation, characteristics and management. *J Res Pharm Pract*. 2014;3(1):12-18. doi:10.4103/2279-042X.132704.

195. Collaborative Group on Hormonal Factors in Breast Cancer CG on HF in B. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*. 2012;13(11):1141-1151. doi:10.1016/S1470-2045(12)70425-4.

196. Micheli A, Muti P, Secreto G, et al. Endogenous sex hormones and subsequent breast cancer in premenopausal women. *Int J Cancer*. 2004;112(2):312-318. doi:10.1002/ijc.20403.

197. Berrino F, Muti P, Micheli A, et al. Serum Sex Hormone Levels After Menopause and Subsequent Breast Cancer. *JNCI J Natl Cancer Inst*. 1996;88(5):291-297. doi:10.1093/jnci/88.5.291.

198. Eden JA. Menopausal status, adipose tissue, and breast cancer risk: impact of estrogen replacement therapy. *Horm Mol Biol Clin Investig*. 2013;14(2). doi:10.1515/hmbci-2013-0007.

199. Honma N, Hosoi T, Arai T, Takubo K. Estrogen and cancers of the colorectum, breast, and lung in postmenopausal women. *Pathol Int*. 2015;65(9):451-459. doi:10.1111/pin.12326.

200. White AJ, Nichols HB, Bradshaw PT, Sandler DP. Overall and central adiposity and breast cancer risk in the sister study. *Cancer*. 2015;121(20):3700-3708. doi:10.1002/cncr.29552.

201. Lacey J V, Kreimer AR, Buys SS, et al. Breast cancer epidemiology according to recognized breast cancer risk factors in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial Cohort. *BMC Cancer*. 2009;9(1):84. doi:10.1186/1471-2407-9-84.

202. Greenberg AS, Obin MS. Obesity and the role of adipose tissue in inflammation and metabolism. *Am J Clin Nutr*. 2006;83(2):461S-465S.

http://www.ncbi.nlm.nih.gov/pubmed/16470013. Accessed December 9, 2016.

203. DeNardo DG, Coussens LM. Inflammation and breast cancer. Balancing immune response: crosstalk between adaptive and innate immune cells during breast cancer progression. *Breast Cancer Res*. 2007;9(4):212. doi:10.1186/bcr1746.

204. Hotamisligil GS, Arner P, Caro JF, Atkinson RL, Spiegelman BM. Increased adipose tissue expression of tumor necrosis factor-alpha in human obesity and insulin resistance. *J Clin Invest*. 1995;95(5):2409-2415. doi:10.1172/JCI117936.

205. Stephens JM, Lee J, Pilch PF. Tumor necrosis factor-alpha-induced insulin resistance in 3T3-L1 adipocytes is accompanied by a loss of insulin receptor substrate-1 and GLUT4 expression without a loss of insulin receptor-mediated signal transduction. *J Biol Chem*. 1997;272(2):971-976. http://www.ncbi.nlm.nih.gov/pubmed/8995390. Accessed December 9, 2016.

206. American Cancer Society. *Breast Cancer Facts & Figures 2013-2014*. Atlanta; 2013.

207. Rose DP, Komninou D, Stephenson GD. Obesity, adipocytokines, and insulin resistance in breast cancer. *Obes Rev*. 2004;5(3):153-165. doi:10.1111/j.1467-789X.2004.00142.x.

208. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer. *Lancet (London, England)*. 1997;350(9084):1047-1059. http://www.ncbi.nlm.nih.gov/pubmed/10213546. Accessed November 10, 2016.

209. Rose DP, Vona-Davis L. Interaction between menopausal status and obesity in affecting breast cancer risk. *Maturitas*. 2010;66(1):33-38. doi:10.1016/j.maturitas.2010.01.019.

210. Hilakivi-Clarke L, Forsén T, Eriksson JG, et al. Tallness and overweight during childhood have opposing effects on breast cancer risk. *Br J Cancer*. 2001;85(11):1680-1684. doi:10.1054/bjoc.2001.2109.

211. Kossman DA, Williams NI, Domchek SM, Kurzer MS, Stopfer JE, Schmitz KH. Exercise lowers estrogen and progesterone levels in premenopausal women at high risk of breast cancer. *J Appl Physiol*. 2011;111(6):1687-1693. doi:10.1152/japplphysiol.00319.2011.

212. Choudhury F, Bernstein L, Hodis HN, Stanczyk FZ, Mack WJ. Physical activity and sex hormone levels in estradiol- and placebo-treated postmenopausal women. *Menopause*. 2011;18(10):1079-1086. doi:10.1097/gme.0b013e318215f7bd.

213. Tsai SA, Stefanick ML, Stafford RS. Trends in menopausal hormone therapy use of US office-based physicians, 2000-2009. *Menopause*. 2011;18(4):385-392.

doi:10.1097/gme.0b013e3181f43404.

214. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321-333. http://www.ncbi.nlm.nih.gov/pubmed/12117397. Accessed November 9, 2016.

215. Beral V, Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet (London, England)*. 2003;362(9382):419-427. http://www.ncbi.nlm.nih.gov/pubmed/12927427. Accessed November 10, 2016.

216. Shapiro S, Farmer RDT, Stevenson JC, Burger HG, Mueck AO, Gompel A. Does hormone replacement therapy (HRT) cause breast cancer? An application of causal principles to three studies. *J Fam Plan Reprod Heal Care*. 2013;39(2):80-88. doi:10.1136/jfprhc-2012-100508.

217. Anderson GL, Limacher M, Assaf AR, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA*. 2004;291(14):1701-1712. doi:10.1001/jama.291.14.1701.

218. Kapil U, Bhadoria AS, Sareen N, Singh P, Dwivedi SN. Reproductive factors and risk of breast cancer: A Review. *Indian J Cancer*. 2014;51(4):571-576. doi:10.4103/0019-509X.175345.

219. Kobayashi S, Sugiura H, Ando Y, et al. Reproductive history and breast cancer risk. *Breast Cancer*. 2012;19(4):302-308. doi:10.1007/s12282-012-0384-8.

220. Ma H, Bernstein L, Pike MC, Ursin G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res*. 2006;8(4):R43. doi:10.1186/bcr1525.

221. Bernstein L, Pike MC, Ross RK, Judd HL, Brown JB, Henderson BE. Estrogen and sex hormone-binding globulin levels in nulliparous and parous women. *J Natl Cancer Inst*. 1985;74(4):741-745. http://www.ncbi.nlm.nih.gov/pubmed/3857369. Accessed November 8, 2016.

222. Shohat-Tal A, Sen A, Barad DH, Kushnir V, Gleicher N. Genetics of androgen metabolism in women with infertility and hypoandrogenism. *Nat Rev Endocrinol*. 2015;11(7):429-441. doi:10.1038/nrendo.2015.64.

223. Mehrgou A, Akouchekian M. The importance of BRCA1 and BRCA2 genes mutations in breast cancer development. *Med J Islam Repub Iran*. 2016;30:369. http://www.ncbi.nlm.nih.gov/pubmed/27493913. Accessed November 8, 2016.

224. Anderson DE, Badzioch MD. Familial breast cancer risks. Effects of prostate and other cancers. *Cancer*. 1993;72(1):114-119.

http://www.ncbi.nlm.nih.gov/pubmed/8508396. Accessed October 14, 2016.

225. Pan H, He Z, Ling L, et al. Reproductive factors and breast cancer risk among BRCA1 or BRCA2 mutation carriers: Results from ten studies. *Cancer Epidemiol*. 2014;38(1):1-8. doi:10.1016/j.canep.2013.11.004.

226. Lichtenstein P, Holm N V., Verkasalo PK, et al. Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343(2):78-85. doi:10.1056/NEJM200007133430201.

227. Rudolph A, Chang-Claude J, Schmidt MK. Gene–environment interaction and risk of breast cancer. *Br J Cancer*. 2016;114(2):125-133. doi:10.1038/bjc.2015.439.

228. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ CK (eds). *Cancer Statistics Review, 1975-2013 - SEER Statistics*. Bethesda, MD; 2016. http://seer.cancer.gov/csr/1975_2013/.

229. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype. *Cancer*. 2007;109(9):1721-1728. doi:10.1002/cncr.22618.

230. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes Control*. 2001;12(8):703-711. http://www.ncbi.nlm.nih.gov/pubmed/11562110. Accessed November 8, 2016.

231. Liu L, Deapen D, Bernstein L. Socioeconomic status and cancers of the female breast and reproductive organs: a comparison across racial/ethnic populations in Los Angeles County, California (United States). *Cancer Causes Control*. 1998;9(4):369-380. http://www.ncbi.nlm.nih.gov/pubmed/9794168. Accessed November 8, 2016.

232. Faggiano F, Zanetti R, Costa G. Cancer risk and social inequalities in Italy. *J Epidemiol Community Health*. 1994;48(5):447-452. http://www.ncbi.nlm.nih.gov/pubmed/7964353. Accessed November 8, 2016.

233. Heck KE, Pamuk ER. Explaining the relation between education and postmenopausal breast cancer. *Am J Epidemiol*. 1997;145(4):366-372. http://www.ncbi.nlm.nih.gov/pubmed/9054241. Accessed November 8, 2016.

234. Carter CL, Jones DY, Schatzkin A, Brinton LA. A prospective study of reproductive, familial and socioeconomic risk factors for breast cancer using NHANES I data. *Public Health Rep*. 104(1):45-50. http://www.ncbi.nlm.nih.gov/pubmed/2493661. Accessed November 8, 2016.

235. Mackillop WJ, Zhang-Salomons J, Boyd CJ, Groome PA. Associations between community income and cancer incidence in Canada and the United States. *Cancer*. 2000;89(4):901-912. http://www.ncbi.nlm.nih.gov/pubmed/10951356. Accessed November 8, 2016.

236. Williams RR, Stegens NL, Goldsmith JR. Associations of cancer site and type with occupation and industry from the Third National Cancer Survey Interview. *J Natl Cancer Inst*. 1977;59(4):1147-1185. http://www.ncbi.nlm.nih.gov/pubmed/903993. Accessed November 8, 2016.

237. Singh GK, Williams SD, Siahpush M, Mulhollen A. Socioeconomic, Rural-Urban, and Racial Inequalities in US Cancer Mortality: Part I—All Cancers and Lung Cancer and Part II—Colorectal, Prostate, Breast, and Cervical Cancers. *J Cancer Epidemiol*. 2011;2011:1-27. doi:10.1155/2011/107497.

238. Risser DR. Cancer incidence and mortality in urban versus rural areas of Texas, 1980 through 1985. *Tex Med*. 1996;92(1):58-61. http://www.ncbi.nlm.nih.gov/pubmed/8599169. Accessed November 8, 2016.

239. Nguyen-Pham S, Leung J, McLaughlin D. Disparities in breast cancer stage at diagnosis in urban and rural adult women: a systematic review and meta-analysis. *Ann Epidemiol*. 2014;24(3):228-235. doi:10.1016/j.annepidem.2013.12.002.

240. Macacu A, Autier P, Boniol M, Boyle P. Active and passive smoking and risk of breast cancer: a meta-analysis. *Breast Cancer Res Treat*. 2015;154(2):213-224. doi:10.1007/s10549-015-3628-4.

241. Chen C, Huang Y-B, Liu X-O, et al. Active and passive smoking with breast cancer risk for Chinese females: a systematic review and meta-analysis. *Chin J Cancer*. 2014;33(6):306-316. doi:10.5732/cjc.013.10248.

242. Reynolds P. Smoking and Breast Cancer. *J Mammary Gland Biol Neoplasia*. 2013;18(1):15-23. doi:10.1007/s10911-012-9269-x.

243. MacMahon B, Trichopoulos D, Cole P, Brown J. Cigarette Smoking and Urinary Estrogens. *N Engl J Med*. 1982;307(17):1062-1065. doi:10.1056/NEJM198210213071707.

244. Gu F, Caporaso NE, Schairer C, et al. Urinary concentrations of estrogens and estrogen metabolites and smoking in caucasian women. *Cancer Epidemiol Biomarkers Prev*. 2013;22(1):58-68. doi:10.1158/1055-9965.EPI-12-0909.

245. Stampfer MJ, Hu FB, Manson JE, Rimm EB, Willett WC. Primary Prevention of Coronary Heart Disease in Women through Diet and Lifestyle. *N Engl J Med*. 2000;343(1):16-22. doi:10.1056/NEJM200007063430103.

246. McKenzie F, Biessy C, Ferrari P, et al. Healthy Lifestyle and Risk of Cancer in the European Prospective Investigation Into Cancer and Nutrition Cohort Study.

Medicine (Baltimore). 2016;95(16):e2850. doi:10.1097/MD.0000000000002850.

247. Harris HR, Bergkvist L, Wolk A. Adherence to the World Cancer Research Fund/American Institute for Cancer Research recommendations and breast cancer risk. *Int J Cancer*. 2016;138(11):2657-2664. doi:10.1002/ijc.30015.

248. Romaguera D, Gracia-Lavedan E, Molinuevo A, et al. Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case-control study. *Int J Cancer*. 2017;141(1):83-93. doi:10.1002/ijc.30722.

249. Nomura SJO, Inoue-Choi M, Lazovich D, Robien K. WCRF/AICR recommendation adherence and breast cancer incidence among postmenopausal women with and without non-modifiable risk factors. *Int J Cancer*. 2016;138(11):2602-2615. doi:10.1002/ijc.29994.

250. McKenzie F, Ellison-Loschmann L, Jeffreys M, Firestone R, Pearce N, Romieu I. Healthy lifestyle and risk of breast cancer for indigenous and non-indigenous women in New Zealand: a case control study. *BMC Cancer*. 2014;14(1):12. doi:10.1186/1471-2407-14-12.

251. Bowman SA, Friday JE, Moshfegh AJ. MyPyramid Equivalents Database, 2.0 for USDA Survey Foods, 2003-2004: Documentation and User Guide. 2008. http://www.ars.usda.gov/ba/bhnrc/fsrg. Accessed May 10, 2017.

252. Hayes RB, Sigurdson A, Moore L, et al. Methods for etiologic and early marker investigations in the PLCO trial. *Mutat Res*. 2005;592(1-2):147-154. doi:10.1016/j.mrfmmm.2005.06.013.

253. Nichols HB, Baird DD, DeRoo LA, Kissling GE, Sandler DP. Tubal ligation in relation to menopausal symptoms and breast cancer risk. *Br J Cancer*. 2013;109(5):1291-1295. doi:10.1038/bjc.2013.433.

254. D'Aloisio AA, Nichols HB, Hodgson ME, Deming-Halverson SL, Sandler DP. Validity of self-reported breast cancer characteristics in a nationwide cohort of women with a family history of breast cancer. *BMC Cancer*. 2017;17(1):692. doi:10.1186/s12885-017-3686-6.

255. Ruan X, Seeger H, Wallwiener D, Huober J, Mueck AO. The ratio of the estradiol metabolites 2-hydroxyestrone (2-OHE1) and 16α-hydroxyestrone (16-OHE1) may predict breast cancer risk in postmenopausal but not in premenopausal women: two case–control studies. *Arch Gynecol Obstet*. 2015;291(5):1141-1146. doi:10.1007/s00404-014-3512-1.

256. DiBello JR, Kraft P, McGarvey ST, Goldberg R, Campos H, Baylin A. Comparison of 3 Methods for Identifying Dietary Patterns Associated With Risk of Disease. *Am J Epidemiol*. 2008;168(12):1433-1443. doi:10.1093/aje/kwn274.

257. Heidemann C, Hoffmann K, Spranger J, et al. A dietary pattern protective against type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)—Potsdam Study cohort. *Diabetologia*. 2005;48(6):1126-1134. doi:10.1007/s00125-005-1743-1.

258. Wold S. PLS for Multivariate Linear Modelling, QSAR: Chemometricmetods in Molecular Design. *Methods Princ Med Chem Van*. 1994.

259. Lunn M, McNeil D. Applying Cox Regression to Competing Risks. *Biometrics*. 1995;51(2):524. doi:10.2307/2532940.

260. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*. 2011;22(1):42-52. doi:10.1097/EDE.0b013e3181f74493.

261. García-Closas M, Lubin JH. Power and sample size calculations in case-control studies of gene-environment interactions: comments on different approaches. *Am J Epidemiol*. 1999;149(8):689-692. http://www.ncbi.nlm.nih.gov/pubmed/10206617. Accessed February 14, 2017.

262. Falk RT, Xu X, Keefer L, Veenstra TD, Ziegler RG. A liquid chromatography-mass spectrometry method for the simultaneous measurement of 15 urinary estrogens and estrogen metabolites: assay reproducibility and interindividual variability. *Cancer Epidemiol Biomarkers Prev*. 2008;17(12):3411-3418. http://www.ncbi.nlm.nih.gov/pubmed/19064556. Accessed December 8, 2016.

263. Tippett, Katherine S, Yasmin S, Cypel. Design and Operation: The Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, 1994–96. *US Dep Agric Agric Res Serv Nationwide Food Surv Rep*. 1997:96-96. https://www.ars.usda.gov/ARSUserFiles/80400530/pdf/Dor9496.pdf. Accessed July 7, 2017.

264. Healy D., Trounson A., Andersen A. Female infertility: causes and treatment. *Lancet*. 1994;343(8912):1539-1544. doi:10.1016/S0140-6736(94)92941-6.

265. Kwa M, Plottel CS, Blaser MJ, Adams S. The Intestinal Microbiome and Estrogen Receptor–Positive Female Breast Cancer. *JNCI J Natl Cancer Inst*. 2016;108(8). doi:10.1093/jnci/djw029.

266. Albenberg LG, Wu GD. Diet and the Intestinal Microbiome: Associations, Functions, and Implications for Health and Disease. *Gastroenterology*. 2014;146(6):1564-1572. doi:10.1053/j.gastro.2014.01.058.

267. Shapira I, Sultan K, Lee A, Taioli E. Evolving Concepts: How Diet and the Intestinal Microbiome Act as Modulators of Breast Malignancy. *ISRN Oncol*. 2013;2013:1-10. doi:10.1155/2013/693920.

268. Plottel CS, Blaser MJ. Microbiome and Malignancy. *Cell Host Microbe*.

2011;10(4):324-335. doi:10.1016/j.chom.2011.10.003.

269. Järvenpää P, Kosunen T, Fotsis T, Adlercreutz H. In vitro metabolism of estrogens by isolated intestinal micro-organisms and by human faecal microflora. *J Steroid Biochem*. 1980;13(3):345-349. http://www.ncbi.nlm.nih.gov/pubmed/7392610. Accessed July 19, 2017.

270. Fuhrman BJ, Feigelson HS, Flores R, et al. Associations of the Fecal Microbiome With Urinary Estrogens and Estrogen Metabolites in Postmenopausal Women. *J Clin Endocrinol Metab*. 2014;99(12):4632-4640. doi:10.1210/jc.2014-2222.

271. Wu GD, Chen J, Hoffmann C, et al. Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science (80-)*. 2011;334(6052):105-108. doi:10.1126/science.1208344.

272. Montel M-C, Buchin S, Mallet A, et al. Traditional cheeses: Rich and diverse microbiota with associated benefits. *Int J Food Microbiol*. 2014;177:136-154. doi:10.1016/j.ijfoodmicro.2014.02.019.

273. Uyeno Y, Sekiguchi Y, Kamagata Y. Impact of consumption of probiotic lactobacilli-containing yogurt on microbial composition in human feces. *Int J Food Microbiol*. 2008;122(1-2):16-22. doi:10.1016/j.ijfoodmicro.2007.11.042.

274. Bøhn SK, Blomhoff R, Paur I. Coffee and cancer risk, epidemiological evidence, and molecular mechanisms. *Mol Nutr Food Res*. 2014;58(5):915-930. doi:10.1002/mnfr.201300526.

275. Thorburn AN, Macia L, Mackay CR. Diet, Metabolites, and "Western-Lifestyle" Inflammatory Diseases. *Immunity*. 2014;40(6):833-842. doi:10.1016/j.immuni.2014.05.014.

276. Grivennikov SI, Greten FR, Karin M. Immunity, Inflammation, and Cancer. *Cell*. 2010;140(6):883-899. doi:10.1016/j.cell.2010.01.025.

277. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin*. 2017;67(1):7-30. doi:10.3322/caac.21387.

278. Guinter et al. Unpublished work (aim 1).

279. Bowman SA, Clemens JC, Friday JE, Thoerig RC, Moshfegh AJ. Food Patterns Equivalents Database 2011-12: Methodology and User Guide: Methodology and User Guide. 2014. http://www.ars.usda.gov/nea/bhnrc/fsrg. Accessed October 13, 2016.

280. Washington DC, 20201. *2008 Physical Activity Guidelines for Americans*. Washington, D.C.; 2008. www.health.gov/paguidelines. Accessed December 20, 2016.

281. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet (London, England)*. 2001;358(9291):1389-1399. doi:10.1016/S0140-6736(01)06524-2.

282. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359-E386. doi:10.1002/ijc.29210.

283. Guinter et al. Unpublished work (aim 2).

284. Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B*. 1972;34:187-220. doi:10.2307/2985181.

285. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982;69(1):239-241. doi:10.1093/biomet/69.1.239.

286. Gavaler JS, Van Thiel DH. The association between moderate alcoholic beverage consumption and serum estradiol and testosterone levels in normal postmenopausal women: relationship to the literature. *Alcohol Clin Exp Res*. 1992;16(1):87-92. http://www.ncbi.nlm.nih.gov/pubmed/1558307. Accessed August 16, 2017.

287. Lin CJ, DeRoo LA, Jacobs SR, Sandler DP. Accuracy and reliability of self-reported weight and height in the Sister Study. *Public Health Nutr*. 2012;15(6):989-999. doi:10.1017/S1368980011003193.

288. Mishra GD, McNaughton SA, Bramwell GD, Wadsworth MEJ. Longitudinal changes in dietary patterns during adult life. *Br J Nutr*. 2006;96(4):735-744. http://www.ncbi.nlm.nih.gov/pubmed/17010234. Accessed October 26, 2017.

289. Tabung FK, Steck SE, Liese AD, et al. Patterns of change over time and history of the inflammatory potential of diet and risk of breast cancer among postmenopausal women. *Breast Cancer Res Treat*. 2016;159(1):139-149. doi:10.1007/s10549-016-3925-6.

290. Newby PK, Weismayer C, Akesson A, Tucker KL, Wolk A. Long-term stability of food patterns identified by use of factor analysis among Swedish women. *J Nutr*. 2006;136(3):626-633. http://www.ncbi.nlm.nih.gov/pubmed/16484535. Accessed October 26, 2017.

291. Weismayer C, Anderson JG, Wolk A. Changes in the stability of dietary patterns in a study of middle-aged Swedish women. *J Nutr*. 2006;136(6):1582-1587. http://www.ncbi.nlm.nih.gov/pubmed/16702325. Accessed October 26, 2017.

292. O'Connor M-F, Irwin MR. Links Between Behavioral Factors and Inflammation. *Clin Pharmacol Ther*. 2010;87(4):479-482. doi:10.1038/clpt.2009.255.

293. Schatzkin A, Kipnis V, Carroll RJ, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. *Int J Epidemiol*. 2003;32(6):1054-1062. http://www.ncbi.nlm.nih.gov/pubmed/14681273. Accessed October 26, 2017.

294. Freedman LS, Commins JM, Moler JE, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. *Am J Epidemiol*. 2014;180(2):172-188. doi:10.1093/aje/kwu116.

295. Tabung FK, Smith-Warner SA, Chavarro JE, et al. Development and Validation of an Empirical Dietary Inflammatory Index. *J Nutr*. 2016;146(8):1560-1570. doi:10.3945/jn.115.228718.