Risk Factors for Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma Incidence in Postmenopausal Women: a Women’s Health Initiative (WHI) Study

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Kati S. Maharry, MAS
Graduate Program in Public Health

The Ohio State University

2016

Dissertation Committee:
Electra D. Paskett, PhD, Advisor
Stanley A. Lemeshow, PhD, Co-Advisor
Peter G. Shields, MD
Rebecca D. Jackson, MD
Copyright by

Kati S. Maharry

2016
Abstract

Background

Various exposures have been investigated by epidemiologic studies as risk factors for leukemia incidence. However, studies focusing on and, therefore, findings particular to Small Lymphocytic Lymphoma (SLL) and Chronic Lymphocytic Leukemia (CLL) have been very scarce and findings, across these studies, have been inconsistent. In fact, according to the NCI, there are only a few established risk factors for CLL/SLL: 1) being middle-aged or older, male, or white; 2) a family history of CLL or cancer of the lymph system; and 3) having relatives who are Russian Jews or Eastern European Jews.\(^1\) As none of these are risk factors that can be altered with lifestyle changes, we sought to explore potential and likely risk factors that can be modified with behavior.

Using the Women’s Health Initiative (WHI), we investigated CLL/SLL risk and its relationship in postmenopausal women with three specific aims of interest: Aim 1) personal habits, comprising diet, drinking habits (including alcohol and coffee, both which have potential biological activity in leukemogenesis), and exercise; Aim 2) hormonal exposures, such as oral contraceptives (OC) and hormone therapies (HT); and Aim 3) pesticide exposures. These three main areas were chosen because there is a great need to understand 1) why CLL/SLL is significantly more prevalent in industrial
countries compared to developing countries, and 2) why men have a two-fold increase in their risk of developing CLL/SLL. Therefore, we hypothesized that CLL/SLL may have a similar etiology to other well established hormone-dependent cancers that are likewise more prevalent in Western countries. For example, it has been shown that the development of hormone-dependent cancers, cancers that are related to sex hormones or sex hormone metabolism, all have a relationship to dietary and other lifestyle factors, which tend to be specific to Western cultures. On the contrary, in certain cultures, such as ones in Asia, where diets are more vegetarian or at least semi-vegetarian and also contain high levels of phytoestrogens (e.g., from soy based foods), the incidence of hormone-dependent cancers is very low. Furthermore, hormone therapy use and the use of oral contraceptives are historically more common in industrialized countries as is the use of pesticides in agriculture and in home lawn care. Of note, in recent decades pesticide application has become more prevalent than it was in the past in the developing world. However, only 25% of the two million tons of pesticides applied worldwide each year is used in the developing world (the rest is applied in Europe and the U.S.).

Methods

Aims 1 and 2

A total of 161,808 postmenopausal women (50–79 years) enrolled in the WHI between September 1, 1993 and December 31, 1998, were followed, on average, for 13.8 years. Conditional logistic regression models were used to assess the effects of personal
habits (Aim 1) and hormonal exposures (Aim 2) on CLL/SLL risk using an age and race matched nested case-control design with 328 confirmed CLL/SLL and 1312 control subjects.

Aim 3

A total of 93,676 postmenopausal women (50–79 years) were enrolled in the WHI Observational Study arm between September 1, 1993 and December 31, 1998, and were followed up for a mean of 13.8 years. Conditional logistic regression models were used to evaluate the effect of pesticide exposure and risk of CLL/SLL using an age and race matched nested case-control design with 175 confirmed CLL/SLL cases and 628 control subjects. Pesticide exposure since the age of 21 was evaluated by questionnaire data at the one-year visit.

Results

Aim 1

After adjustment for potential confounders, coffee drinking showed a weak association with lower CLL/SLL risk, for women enrolled on the clinical trials (CT), but not on the observational study (OS) of the WHI. Women enrolled on the CT who consumed coffee on a regular basis had lower risk of CLL/SLL (odds ratio (OR) = 0.73, 95% confidence interval (CI): 0.51, 1.05; P=.09; CT), compared to non-coffee drinkers. Past oral contraceptive use (OR=0.74, 95% CI: 0.56, 0.96; P=.03; CT+OS) and obesity
(OR=0.71, 95% CI: 0.53, 0.94; P=.20; CT+OS) both showed to be protective against CLL/SLL, whereas past estrogen use (OR=1.32, 95% CI:1.02, 1.71; P=.04; CT+OS) was adverse. Neither region of the U.S. nor smoking were significantly related to CLL/SLL risk. We did not find any significant associations with other personal habits and risk of CLL/SLL, such as alcohol use, dietary factors, or exercise habits. In addition, the ACS Nutrition and Physical Activity Cancer Prevention Guidelines score, a composite score based on diet, BMI, exercise, and alcohol use, did not associate with CLL/SLL risk in our cohort of women.

**Aim 2**

After adjustment for potential confounders, past OC users had lower CLL/SLL risk than women with no past OC exposure (OR= 0.73, 95% CI: 0.56, 0.96; P=.02). Past hormone therapy (HT) use – specifically estrogen-alone (E-alone) therapy – however, conferred higher CLL/SLL risk (OR = 1.32, 95% CI: 1.01, 1.71; P=.04). CLL/SLL risk was lower for obese than non-obese women (OR=0.73, 95% CI: 0.55, 0.97; P=.03), and obese women who took OC (i.e., women with both high endogenous and exogenous estrogen exposures) had about half the CLL/SLL risk compared to no OC users or to non-obese OC users (P=.01). Current HT did not associate with risk of CLL/SLL (P=.64).

**Aim 3**

After adjustment for potential confounders, women with any pesticide exposure during adulthood (either at work or in their home), had a significantly higher risk for
developing CLL/SLL than did women with no pesticide exposure (OR= 1.65, 95% CI: 1.11, 2.45; P=.01). The significant association between pesticide exposure and CLL/SLL incidence was independent of region of the US, smoking and obesity status.

Conclusions

In the first aim of our study we found that postmenopausal women on the WHI CT who had a habit of regular daily coffee drinking had a reduced risk for CLL/SLL, by 27%. Coffee is composed of polyphenols and caffeine containing many bioactive compounds that can potentially contribute to the protective association we identified. One such biological theory in support of our results is that coffee has been shown to be mildly estrogenic. Women who consume high amounts of coffee excrete more estrogen in their urine compared to those who drink less. Phytoestrogens identified in coffee potentially contribute to this, in particular it has been shown that blood enterolactone levels increase due to coffee drinking and likely exhibit weak estrogenic effects. This mechanism could lead to reduction in the risk of CLL/SLL through pathways we postulated in our study of hormones and risk of CLL/SLL (Aim 2). Other drinking, dietary, or exercise habits had no impact on CLL/SLL incidence, neither by specific subcomponents, nor by the ACS score. These null results are in concordance with the overall lack of established dietary and physical activity risk factors for CLL/SLL.

In the second aim of our study we showed that postmenopausal women with past - but not current- HT use (E-alone in particular) had an increased CLL/SLL risk, whereas OC use and obesity were protective. We concluded from this investigation that very high
levels of circulating estrogen in a woman’s body from either endogenous (such as a result of obesity) or exogenous sources (as a result of OC use), or both, likely offer protection against CLL/SLL. Furthermore, this protection due to estrogen can also explain the lower incidence of CLL/SLL in women relative to men. On the contrary, E-alone therapy, mainly prescribed to women undergoing bilateral oophorectomy (BOO), increased the risk of CLL/SLL possibly due to lower levels of endogenous estrogen despite exogenous therapy.

In the third aim, we found that women with a history of pesticide exposure had an increased risk of CLL/SLL, irrespective of the location (work or home) of the exposure. Although specific information regarding the type of pesticide chemicals used was not collected in our study, we know that our cohort of women would have been exposed to around 100 different chemicals, many of which are proven carcinogens with known serious health effects. Certain pesticides have known endocrine disturbing properties, such as estrogenic activity, however, their DNA damage potential, which can lead to chromosomal alterations, is likely to be a more dominant factor with respect to carcinogenesis. Therefore, we postulate that the adverse impact we and others have observed relative to pesticides is dominated by the DNA damaging activity rather than estrogenic activity.

Overall, we can conclude that the risk factors that seem to matter the most for CLL/SLL risk, such as coffee drinking, obesity, and OC use, all have estrogenic effects that, in this case, appear to offer protection. The biological mechanisms leading to these findings would need to be further investigated to understand exact pathways. Although certain pesticides have also been known to alter estrogen in the body, the main impact of
pesticides in our study is most likely due to various carcinogens causing DNA damage.

Our findings consistently identified risk factors mainly specific to industrialized countries that can help explain the higher incidence of CLL/SLL in the Western world compared to that found in developing countries. In addition, we showed a probable association of estrogen dominance being protective against CLL/SLL, which can help in understanding the lower rates of this hematological cancer in women relative to men.
Dedication

Dedicated to my parents, Annamaria and Charles, who inspired me to follow them on a path of higher science education; my children, Aaron, Sophie, and Mia; and my extremely patient husband, Rob.
Acknowledgements

First and foremost, this research would not have been possible without the great guidance and diligence of my advisor, Dr. Electra Paskett. Through this work, I received the most rigorous epidemiology training from Dr. Paskett. I firmly believe that this skillset will allow me to further my career as a cancer researcher, with a greater sense of inquisitiveness and knowledge of research methods. Also, I am very grateful to my biostatistics advisor, Dr. Stanley Lemeshow whose statistical guidance enabled me to apply my skills acquired in graduate school to real life data, while holding me to sound statistical methodology he taught me in class. In addition, I am extremely grateful to the other two members of my esteemed committee: Dr. Rebecca Jackson and Dr. Peter Shields. Dr. Jackson provided expertise regarding the Women’s Health Initiative (WHI), both relative to study structure and to other relevant WHI research, that was of great benefit in advancing my work successfully. Dr. Shields with his vast clinical oncology and cancer epidemiology expertise convinced me to focus my research on a specific type of adult leukemia, one that is the least known about relative to modifiable risk factors. His vision and guidance allowed me to discover novel risk factors in this challenging research area, that distinguish risk for women compared to men.

In addition, I am grateful to my daughter, Sophie, who proofread my writing and with each draft managed to teach me more about the English language, and to Cecilia DeGraffinreid for her continued support in all administrative tasks regarding regulatory
requirements for this research.

I want to also acknowledge all the women who participated in the WHI and contributed to the biggest and best characterized research study of postmenopausal women to date. Without the commitment of the large cohort of the WHI women, we would not have been able to successfully investigate a rare cancer, such as chronic lymphocytic leukemia / small lymphocytic lymphoma.

Lastly, I am forever grateful to Dr. Clara Bloomfield and Dr. Guido Marcucci who have both provided me with the utmost support and mentoring in my pursuit of this doctoral degree. I feel very lucky and honored to have had the opportunity to collaborate with them for many years as part of their world-class acute myeloid leukemia research group. It is their excellence that prompted me to focus my studies on leukemia.
Vita

EDUCATION
9/2010-4/2016 (exp.) The Ohio State University, Columbus, OH
PhD. in Epidemiology (Major: Cancer Epidemiology/Minor: Biostatistics)

6/1996-6/1998 The Ohio State University, Columbus, OH
Masters of Applied Statistics

9/1994-3/1995 The Ohio State University, Columbus, OH
Preventive Medicine Masters of Science Program

9/1990-6/1994 The Ohio State University, Columbus, OH
B.S. in Industrial and Systems Engineering

Publications


Whitman SP, Kohlschmidt J, Maharry K, Volinia S, Mrózek K, Nicolet D, Schwind


Whitman SP, Caligiuri MA, Maharry K, Radmacher MD, Kohlschmidt J, Becker H,


**Fields of Study**

**Major Field:** Public Health  
Epidemiology

**Minor Field:** Biostatistics
Abstract ........................................................................................................................................... ii
Background ....................................................................................................................................... ii
Methods ........................................................................................................................................... iii
  Aims 1 and 2 ................................................................................................................................... iii
  Aim 3 ............................................................................................................................................... iv
Results .............................................................................................................................................. iv
  Aim 1 .............................................................................................................................................. iv
  Aim 2 ............................................................................................................................................... v
  Aim 3 ............................................................................................................................................... v
Conclusions ...................................................................................................................................... vi
Dedication .......................................................................................................................................... ix
Acknowledgements .......................................................................................................................... x
Vita ....................................................................................................................................................... xii
List of Tables ..................................................................................................................................... xxv
List of Figures ..................................................................................................................................... xxvi
Chapter 1: Introduction ....................................................................................................................... 1
Chapter 2: Background ....................................................................................................................... 10
  2.1 Introduction .............................................................................................................................. 10
  2.2 Types of Leukemia .................................................................................................................... 10
  2.3 Recently Revised Guidelines for Chronic Lymphocytic Leukemia Classification .............. 12
  2.4 Risk Factors for Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma .......... 13
    2.4.1 Behavioral/Lifestyle Risk Factors for Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma ......................................................................................................................... 14
      2.4.1.1 Smoking .......................................................................................................................... 14
      2.4.1.2 Alcohol .......................................................................................................................... 15
      2.4.1.3 Obesity / Physical Activity / Diet .................................................................................... 17
      2.4.1.4 Hair Dyes ....................................................................................................................... 21
      2.4.1.5 Nonsteroidal Anti-inflammatory Drugs ....................................................................... 22
      2.4.1.6 Hormone Therapy ......................................................................................................... 24
    2.4.2 Environmental Risk Factors for Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma .......................................................................................................................... 25
      2.4.2.1 Radiation ....................................................................................................................... 26
      2.4.2.2 Secondhand Smoke ....................................................................................................... 28
      2.4.2.3 Pollution ......................................................................................................................... 29
      2.4.2.4 Pesticides and Herbicides as Occupational Exposures .............................................. 31
      2.4.2.5 Benzene as Occupational Exposure ............................................................................. 34
      2.4.2.6 Viruses .......................................................................................................................... 36
Chapter 3: Methods......................................................................................................................... 61
 3.1 Overview.................................................................................................................................. 61
 3.1.1 Extension Studies..................................................................................................................... 65
 3.1.2 Additional Criteria for the Current Study.................................................................................. 65
 3.2 Research Design .......................................................................................................................... 66
 3.3 Data Source ................................................................................................................................ 66
 3.3.1 WHI Database ......................................................................................................................... 66
 3.3.2 Clinical Trials Database ........................................................................................................... 67
 3.3.3 Observational Study Database ............................................................................................... 69
 3.4 Study Period and Population ...................................................................................................... 70
 3.5 Data Elements ............................................................................................................................ 71
 3.5.1 Demographics ......................................................................................................................... 72
 3.5.2 Medical History ....................................................................................................................... 73
 3.5.3 Reproductive History .............................................................................................................. 73
 3.5.4 Personal Habits ........................................................................................................................ 74
 3.5.5 Family History ........................................................................................................................ 77
 3.5.6 Miscellaneous Repeated Measures Variables ....................................................................... 78
 3.5.7 Outcome .................................................................................................................................. 80
 3.6 Statistical Analyses .................................................................................................................... 81
 3.6.1 Sample Size Calculation ......................................................................................................... 81
 3.6.2 Descriptive Analyses .............................................................................................................. 82
 3.6.3 Analysis of Personal Habits as Risk Factors for CLL/SLL (Specific Aim 1)......................... 83
 3.6.4 Analysis of Hormone Exposures as Risk Factor for CLL/SLL (Specific Aim 2)............. 84
 3.6.5 Analysis of Pesticide Exposures as Risk Factors for CLL/SLL (Specific Aim 3)........... 85
 3.6.6 Summary of Statistical Analyses Methods .......................................................................... 86

Chapter 4: Personal Habits and their Association with Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma Incidence ............................................................................. 87
 4.1 Background................................................................................................................................ 87
 4.2 Methods ................................................................................................................................... 92
 4.2.1 Study Design .......................................................................................................................... 92
 4.2.2 Measures ............................................................................................................................... 93
 4.2.3 Statistical analysis ................................................................................................................. 95
 4.3 Results ...................................................................................................................................... 96
 4.3.1 Descriptive Data Analysis ..................................................................................................... 96
 4.3.2 Multivariable Model Evaluating CLL/SLL Risk................................................................. 97
 4.4 Discussion ................................................................................................................................. 98

Chapter 5: Hormone Exposure and Incidence of Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma ........................................................................................................ 112
 5.1 Background .............................................................................................................................. 112
 5.2 Methods .................................................................................................................................. 115
Chapter 6:  Pesticide Exposure and Incidence of Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma

6.1 Background......................................................................................................................... 134
6.2 Methods............................................................................................................................... 136
   Study Design.......................................................................................................................... 136
   Measurements......................................................................................................................... 137
   Statistical analysis................................................................................................................ 139
6.4 Results.................................................................................................................................. 140
6.5 Discussion.............................................................................................................................. 142

Chapter 7.  Summary, Conclusions, Implications................................................................. 151
7.1 Summary.............................................................................................................................. 151
7.2 Conclusions........................................................................................................................ 153
7.3 Limitations, Strengths, and Implications............................................................................ 156

Appendix : Women’s Health Initiative Data Distribution Agreement.......... 160
Bibliography ........................................................................................................................... 163
List of Tables

Table 1.1 Reviewed articles on risk factors for CLL/SLL................................................................. 43
Table 4.1. Baseline Characteristics and WHI Study Participation by Control and Case Status...........102
Table 4.2. Baseline Characteristics by WHI Study Status..................................................................107
Table 4.3 Conditional Logistic Regression Multivariable Modeling for CLL/SLL...............................111
Table 5.1 Baseline Characteristics and WHI Study Participation by Control and Case Status..........128
Table 5.2 Exogeneous Hormone Use by Control and Case Status.........................................................131
Table 5.3 Conditional Logistic Regression Multivariable Modeling for CLL/SLL...............................132
Table 6.1 Baseline Characteristics by Control and Case Status..........................................................147
Table 6.2. Pesticide Exposure by Control and Case Status.................................................................149
Table 6.3 Conditional Logistic Regression Multivariable Model for CLL/SLL......................................150
List of Figures

Figure 1.1 Venn Diagram of the WHI study design ...................................................... 64
Figure 5.1. Estrogen exposures and risk of CLL/SLL ..................................................... 133
Chapter 1: Introduction

Leukemia is a hematological cancer of the bone marrow and/or blood that occurs when there is an abnormal increase of malignant immature white blood cells, or blasts. There are four main types of leukemia: Acute Myeloid Leukemia (AML), Acute Lymphocytic Leukemia (ALL), Chronic Lymphocytic Leukemia (CLL)/Small Lymphocytic Lymphoma (SLL) and Chronic Myeloid Leukemia (CML). Although certain subtypes of leukemia have better prognoses than others, the cause of the disease is unknown in most patients and treatment is generally unsuccessful (5-year survival rates among adults 65 years and older: 5-10% for AML and ALL, 20% for CML, and 60% for CLL/SLL). Although leukemia is a relatively rare form of cancer in adults (approximately 3 % of all cancers), there are about 240,000 new cases diagnosed annually worldwide (43,800 in the U.S.) and approximately 200,000 deaths associated with it (23,300 in the U.S.) each year. Leukemia is ranked fifth in person-years of life lost due to cancer, directly behind breast and pancreatic cancer. In addition, according to the National Cancer Institute’s Surveillance, Epidemiology and End Results (SEER) Program, during recent decades (between 1975 and 2010), there has been an increase among females in the U.S. of developing leukemia.
In industrialized countries CLL/SLL is the most common type of adult leukemia. Despite being a frequent type of leukemia, CLL/SLL is a relatively rare form of cancer (approximately 1% of all cancers). There are still about 15,720 new CLL/SLL cases diagnosed each year in the U.S – mainly in older adults (the lifetime risk of CLL/SLL is 0.52%).\textsuperscript{7,8,10} Among men, the rate of CLL/SLL is twice compared to women and the reasons behind this have not been established. In addition, according to the National Cancer Institute’s SEER Program, there has been an increased incidence of CLL/SLL between 1975 and 2010 among white females in the U.S.\textsuperscript{6} Furthermore, though CLL/SLL is historically extremely rare among Asians, a rising incidence through a birth cohort effect has been noted in recent years. Asian populations born in the U.S. have the most notable increase in incidence, thereby supporting the notion that environmental risk factors are likely contributors to this disease.

Due to the potential asymptomatic nature of CLL/SLL – the disease can go undetected for several years prior to diagnosis - reliable global estimates for CLL/SLL incidence are not available. Still, it is known that the incidence of CLL/SLL varies by geographic region, with the U.S., Europe, and Australia having the highest number of new diagnoses yearly. While certain subtypes of CLL/SLL have better prognoses than others, and patient outcome also depends on the stage of the disease, the causes of the disease are not identified for the majority of patients and on average only about 60% survive past 5 years.\textsuperscript{6}

There have been very few CLL/SLL or other leukemia risk factors confirmed to
date. Accordingly, in most CLL/SLL cases the cause of the disease is unknown. In general, it is thought that most cancers are associated with environmental, genetic, behavioral, or lifestyle risk factors. In a recent analysis of cancer tissues, Tomasetti and Vogelstein showed a statistically significant relationship between the total number of stem cell divisions in tissue and the lifetime risk of cancer in that tissue. Environment and genetics are the two pathways that can influence the number of cell divisions- the authors directly attribute these two risk factors to explain a third of the variation of cancer. The remaining variation they attribute to random mutations during DNA replication- a process not independent of environmental and genetic factors that can be influenced at an individual level during one’s lifetime. In their research, they show that the risk for CLL/SLL is mainly (but not exclusively) attributable to stochastic factors related to DNA replication errors and among cancer types with a similar finding, CLL/SLL ranks in the middle- indicating that in addition to stochastic factors, direct environmental and genetic factors may also play an important role. The extent of this finding for CLL/SLL is yet to be confirmed in a population based setting.

For leukemia overall, established or strongly suspected/hypothesized and potentially addressable risk factors are: ionizing radiation (natural and artificial), previous chemotherapy, family history of cancer (breast for acute leukemia, hematological for CLL/SLL), exposure to benzene, certain plant protection products, hair dyes, obesity, acetaminophen use, viruses (ex. human T-lymphotropic virus), and cigarette smoking. These exposures are thought to lead to leukemia development by causing chromosomal alterations and mutations in the DNA which result in
oncogenesis and/or the deactivation of tumor suppression mechanisms leading to
disruption of the normal lymphoid or myeloid differentiation process. In addition, there
are known inherited genetic factors that account for a small proportion of leukemia cases,
such as certain familial genetic predispositions or chromosomal abnormalities (ex.
Fanconi anemia, Down syndrome and clonal chromosomal mosaicism).\textsuperscript{22,23,24} Moreover,
to date, there are only a few protective factors suspected for adult leukemia overall,
including certain dietary factors, a healthy BMI, daily aspirin use, and light to moderate
beer consumption.\textsuperscript{25,26,27}

Because the aforementioned previously identified risk factors are likely to only
account for a fraction of cases, other exposures remain to be identified. Several studies to
date have investigated risk factors for leukemia, largely focusing on one specific risk
factor at a time and often combining all types of leukemia. The results of these studies
vary, and are often inconclusive or conflicting. Aside from a few post-hoc meta-
analyses, no epidemiologic study to date has investigated a large pool of risk factors
collectively for a specific leukemia subtype, such as CLL/SLL within one cohort of subjects. In
addition, risk factors for CLL/SLL incidence have not been specifically evaluated among a large
cohort of women at the age they are the most prone to develop this type of leukemia. A large
comprehensive study, such as the Women’s Health Initiative (WHI), allows the
evaluation of a wide range of possible risk factors and their complex interactions for
CLL/SLL development among women in the age group most prone to develop this type
of leukemia (median age of CLL/SLL incidence is 70 years).\textsuperscript{21}
Chronic Lymphocytic Leukemia/ SLL is a challenging cancer to detect and subsequently to treat, with numerous, often serious and life threatening adverse events due to drug toxicity. Despite advances in therapy, the resulting long-term outcome remains sub-optimal for the majority of CLL/SLL patients. Current clinical research in CLL/SLL tends to mainly focus on finding optimal therapy (chemotherapy and/or molecularly targeted agents) with consideration of the genetics of the patient (gene mutations, expression, and epigenetics) at diagnosis. By identifying potentially preventable risk factors that lead to CLL/SLL development in the first place, we could prevent a portion of people from ever having to go through often debilitating treatment.

In an effort to focus on CLL/SLL prevention, three manuscripts were developed using well-documented epidemiological data collected from a large, prospective longitudinal study on postmenopausal women. There is a great need to understand 1) why men have a two-fold increase in their risk of developing CLL/SLL and, 2) why CLL/SLL is significantly more prevalent in industrial countries compared to developing countries, and 1) why men have a two-fold increase in their risk of developing CLL/SLL. The WHI offered extensive information on several modifiable risk factors and thus, three aims were selected. The manuscripts sought to address the following aims: Aim 1) personal habits, comprising diet, drinking habits (including alcohol and coffee, both which have potential biological activity in leukemogenesis), and exercise; Aim 2) hormonal exposures, such as oral contraceptives (OC) and hormone therapies (HT); and Aim 3) pesticide exposures.. The WHI offered an opportunity to study the personal habits of women in great detail, such as evaluating
their use of estrogens, specific dietary and drinking habits, as well as their exercise habits. Thus, we were anticipating to discover unique factors that would offer them protection to women against CLL/SLL. By investigating hormonal exposures, we were interested in seeing if estrogen exposure, which is primarily prevalent in women, influences CLL/SLL risk and therefore explains the two-fold risk difference between men and women. In addition, particularly relative to diet, we wanted to explore any habits that are mainly specific to industrialized cultures, such as red meat consumption. Regarding pesticide use history, given that some earlier other studies have indicated an adverse relationship with Non-Hodgkin Lymphomas (NHL) in general, we sought to explore a similar relationship with CLL/SLL. This finding could in turn provide some insight into the increased incidence of CLL/SLL in industrialized countries, where pesticide application is triple in volume, compared to developing countries.

While this is not the first study to investigate risk factors for CLL/SLL in women, to our knowledge it is the largest and best characterized one, focusing solely on women in the age group most at risk of developing CLL/SLL. Relative to the 813 leukemia cases confirmed in the WHI, of which over half (n=450) are CLL/SLL, previous studies that were conducted investigating leukemia incidence among adult women, such as the EPIC Study, the Million Women Study, the Iowa Women’s Health Study, and the Minnesota Cancer Surveillance System Study, all had fewer incident leukemia cases (and therefore fewer CLL/SLL cases) than the WHI (female leukemia cases: n=303 in the EPIC Study, n=428 in the Million Women Study, n=201 in Iowa Women’s Health Study, n=278 in Minnesota Cancer Surveillance System Study).
To address the lack of research pertaining to risk factors specifically for CLL/SLL among women we generated the following specific aims and hypotheses:

**Specific aim 1:** To investigate a set of personal habits factors, such as diet, smoking, drinking habits, and physical exercise and identify their associations with CLL/SLL incidence in the WHI.

**Hypothesis 1:** Women with healthier personal habits have a decreased incidence of CLL/SLL compared to women with less healthy personal habits, after controlling for other potentially confounding factors, such as weight. We hypothesize that healthy lifestyles such as more fruit and vegetable consumption, lower fat diet, less red/smoked meat in the diet, no smoking, and more exercise protects women against CLL/SLL through various mechanisms, similar to what has been found in some other cancers.

**Specific aim 2:** To explore the use of hormone therapy as a risk factor for CLL/SLL development among women in the WHI, by also considering other risk factors, such as oral contraceptive use, smoking, and obesity.

**Hypothesis 2: 2.a)** Women who have used HT in the past (estrogen only, or estrogen and progesterone combination) have a higher incidence of CLL/SLL, due to estrogen being a potential promoter of cancer cells (through estrogen metabolism-mediated oxidative stress), compared to women who have not used HT in the past, and that HT use remains a significant risk factor for CLL/SLL even after controlling for other important risk factors. **2.b)** Obese women (BMI > 30 after menopause) who have used HT in the past have a higher incidence of
CLL/SLL, due to fat cells storing more estrogen and estrogen being a potential promoter of cancer cells, compared to women who are not obese, irrespective of HT use.

**Specific aim 3:** To investigate the use of various pesticides, herbicides, and insecticides as risk factors associated with CLL/SLL development, in the WHI, a large population-based study among women in the U.S., at the age they are most likely to develop this type of cancer.

**Hypothesis 3:** 3.a) Women with a history of living and/or working in environments with increased use of pesticides, herbicides, and insecticides have a higher incidence of CLL/SLL, compared to those with a history of lower exposures. 3.b) Women who have a history of pesticide exposure at work (i.e., farming occupation) have higher incidence of CLL/SLL compared to those who were not exposed to pesticides at work. 3.c) Women who have a history of applying pesticides themselves at home have higher incidence of CLL/SLL compared to those who have a history commercial lawn service use or to those without history of use pesticides at home, irrespective of occupational pesticide exposure.

In order to explore these research aims, data collected and compiled by the WHI study were analyzed in a nested case-control study design setting. The WHI study was intended to explore the most frequently occurring causes of morbidity and mortality in postmenopausal women, including cancer. As part of this study, detailed information regarding diagnoses of various cancer types was obtained, including subtypes of leukemia- with CLL/SLL as one of the categories. Additionally, a wide range of variables
were collected as part of the WHI study at baseline, as well as at follow-up time points, which were evaluated for risk factor analyses in the current study.
Chapter 2: Background

2.1 Introduction

While Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma (CLL/SLL) is the most common adult leukemia, it still only accounts for a small fraction (1%) of all cancer diagnoses. However, it is a cancer that has been increasing in incidence among women in industrialized countries during recent decades. Although some of this rise in CLL/SLL is likely attributable to an aging population in the developed world, environmental, behavioral, and biological risk factors can all have potential effects on this increasing trend. Below, we consider various conceivable risk factors for CLL/SLL. Whereas a few risk factors already appear to have a stronger hypothesis of being associated with CLL, most of these exposures have only been studied for leukemia overall and not by sub-types, such as CLL/SLL.

2.2 Types of Leukemia

Leukemia is a group of heterogeneous hematological neoplasms of the white blood cells in bone marrow and/or blood. There are two main varieties of leukemia based on lineage, myeloid and lymphoid. Both of these types can be either acute or chronic with
respect to the timing of their clinical presentation. In acute leukemia, the bone marrow cells cannot mature properly, and the disease progresses rapidly - patients would die within a few months without treatment. In chronic leukemia, the bone marrow cells can mature only partly, which leads to a more slowly progressing disease that patients can live with for many years without treatment - often without recognizing for a prolonged period that they are afflicted by the disease.

According to their origin (i.e., lineage) and clinical presentation (acute vs. chronic), there are four main types of leukemia that account for about 85% of all diagnoses with this bone marrow cancer: AML, ALL, CML, CLL/SLL. The remaining 15% consists of a mixture of much more rare subtypes. Leukemia can occur in both pediatric and adult populations. While among children, leukemia – mostly ALL – is the most common type of cancer, it is a relatively rare form of cancer in adults. The most common types of leukemia in adults are AML, CLL/SLL, and CML.

In acute leukemia, AML and ALL, the number of malignant blasts increases rapidly and the disease progresses fast, requiring immediate treatment. AML is a neoplastic disorder of the hematopoietic precursor cells of the bone marrow, whereby the bone marrow is gradually replaced by blast cells. ALL is a neoplastic disorder of the lymphopoietic precursor cells in the bone marrow. In ALL, progressive medullary and extramedullary accumulations of lymphoblasts are present that lack the potential for differentiation and maturation.
In chronic leukemia, CML and CLL/SLL, blast cell counts increase less rapidly than in acute leukemia. Therefore, chronic leukemia gets worse more gradually and sometimes does not require treatment for several years. In CML, there is an uncontrolled proliferation of granulocytes, and usually erythroid cells and megakaryocytes. In CLL/SLL, there is a monoclonal expansion of lymphocytes whereby small B-cell lymphoid neoplasms are composed of monoclonal memory B cells - typically expressing CD23 and the T-cell-associated antigen CD5.

As mentioned earlier, among the four main types of leukemia, CLL/SLL is the most frequently occurring one among adults and - despite still being more common in men- it is also the leukemia that has shown to have an increased incidence among women in recent decades. This recent trend could be potentially attributed to common modern-day exposures, for example hormone therapy or the increased use of pesticides that cause potentially carcinogenic compounds to accumulate in tissue and contribute to DNA damage during one’s lifetime. Therefore, we decided to focus our research on investigating risk factors specifically for CLL/SLL development with the hypothesis of identifying preventable measures against leukemogenesis in older age.

2.3 Recently Revised Guidelines for Chronic Lymphocytic Leukemia Classification

According to the World Health Organization (WHO) from 2008, small lymphocytic lymphoma (SLL) and CLL are to be considered one disease and one entity for disease classification as malignant cells in both diagnoses exhibit the same
immunophenotype. Since its publication, this new definition of CLL has been widely adopted in the clinical setting. On the contrary, with respect to clinical and epidemiologic research, the old categorization – which excludes SLL - remains to be extensively used, particularly in publications. This is most likely due to the frequently extensive lag time involved in the publication process.

In our study, we followed the most current guidelines (WHO 2008) that include SLL cases as CLL, and therefore all references to CLL within our study include both CLL and SLL cases (i.e., CLL/SLL). By correctly including SLL cases in the CLL disease category, the sample size of CLL cases in our study increases by about a third from what it would have been had we used the old classification. This larger sample size allows for a more accurate assessment of risk factors and for a more correct categorization of CLL cases, compared to the old classification.

2.4 Risk Factors for Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma

It is suspected that while a small portion of leukemia cases are genetic, most cases of leukemia result from damage to one’s genes during their lifetime – similarly to what has been found in other forms of cancers. It is therefore important to recognize the different exposures that lead to CLL/SLL, especially since this is almost exclusively a cancer of older people who have been subjected to several decades of DNA damage. These exposures can be categorized into three different groups: 1) behavioral or lifestyle, 2) environmental, and 3) biological risk factors. Behavioral risk factors are potentially
highly modifiable, as are certain environmental ones. On the other hand, biological ones, especially hereditary factors, are less modifiable. By understanding the interplay of genetics and behavioral/environmental exposures, however, there is potential for decreasing one’s risk of developing leukemia despite genetic disposition.

2.4.1 Behavioral/Lifestyle Risk Factors for Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma

Several health behavior and lifestyle risk factors have been investigated for their association with CLL/SLL incidence. However, there has been very little shown to date with respect to their consistent impact on this disease. The following behavioral risk factors have so far been investigated in larger studies for leukemia incidence in general, and in a few instances for CLL specifically: smoking, alcohol use, diet, physical activity, hair dye use, nonsteroidal anti-inflammatory drug use, and hormone therapy. As these are risk factors people choose to be exposed to, their exposure can potentially be altered by behavior modification.

2.4.1.1 Smoking

Despite the association of cigarette smoking with an increased risk for AML, ALL, and CML, no adverse associations with CLL have been found. Moreover, a recent study by Slager et al. showed a modest protective effect of smoking for CLL. This finding has not been so far supported by any biological mechanism and is contrary to research showing increasing biological evidence for the mechanism between smoking and
leukemia incidence for AML, ALL, and CML.\textsuperscript{36} There are over 7,000 chemicals in cigarettes, about 70 of which are known carcinogens.\textsuperscript{37} This list includes toxic agents such as benzene (discussed in more detail in section 2.3.2.6), formaldehyde, polonium 210, arsenic, lead, and ammonia- with benzene having the strongest association with leukemia.\textsuperscript{38} In a recent laboratory study, Zhu et al. showed that hydroquinone, the major metabolite of benzene in humans, increases apoptosis of human bone marrow hematopoietic stem cells, induces DNA double strand breaks in human bone marrow hematopoietic stem cells, and decreases stem cell differentiation and proliferation.\textsuperscript{39}

2.4.1.2 Alcohol

Alcohol consumption and its relationship to leukemia in general has been studied by multiple investigators, with varying conclusions. Some studies point to a protective effect, whereas others show an adverse or no effect on leukemia development. There is a large degree of difficulty in consistently evaluating alcohol consumption: specifically there are complications with reliably measuring frequency, amount, and type of alcohol. Also, controversy exists among researchers as to the reliability of self-reported data in this context, which makes examining this exposure relatively complex.

Although the majority of studies investigating this relationship have not identified any significant association between alcohol use and leukemia incidence to date, there are a few that have. While none of the studies to date have specifically focused on CLL, a few have included sub-analyses for CLL. In a multi-center case-control study done in
Italy, Gorini et al. in 2006 saw a possible J shaped dose-response relationship between alcohol consumption and CLL/SLL incidence. They attributed their findings to the hypothesis that moderate alcohol consumption may have a protective effect due to some specific mechanisms, mainly from antioxidants (resveratrol in wine and flavonoids in beer) and from improved cellular and humoral immune responses. On the other hand, heavy drinking can lead to an impaired immune system, facilitating leukemia development.

A recent study was conducted in Sweden among subjects with alcohol-use disorders, and found a reduced incidence in this population of leukemia in general (as well as for the subset of CLL cases), also supporting a possible protective mechanism despite the large alcohol dose exposure. While this study confirmed the protective effect found in the earlier studies, it was conducted in a unique population of subjects with heavy alcohol consumption, therefore making their results potentially inapplicable to the general population. Conversely, another recent, large prospective study conducted in Northern Europe, failed to show a protective effect and instead suggests a possible positive relationship between alcohol (any type) consumption and the incidence of leukemia, specifically CLL.

The above findings however, may not be contradictory as both the nutrient composition and dose of alcohol are likely to play a simultaneously important role in the mechanism of leukemogenesis. A review article by Diaz et al. provides support for moderate alcohol consumption to improve immune response but for heavy alcohol use to
impair immune function, which may explain at least some of the variation in these findings and sheds light on the complexity of the biological mechanism and the need for accurate assessment of the quantity and quality of alcohol intake.  

2.4.1.3 Obesity / Physical Activity / Diet

In the U.S., it is estimated that over one third of cancer cases can be prevented by improving lifestyle factors related to poor diet, physical inactivity, and obesity. Yet, the percentage of leukemia cases that might be attributed to these lifestyle factors is unknown, and subsequently the percentage breakdown for CLL/SLL, specifically. Although, according to the NCI, leukemia is not included in the list of cancers that have already been identified to be definitively associated with obesity, obesity may increase the risk of developing leukemia, as shown by some studies. The NCI calls for additional studies to be done to confirm or disprove these associations with leukemia incidence. There appears to be a shortage of adequately powered studies investigating obesity and the association of dietary factors and exercise with the development of leukemia.

Obesity is a major public health problem and a serious condition as it raises the risk of a multitude of chronic diseases, including cancer. Several biological mechanisms, focusing on the roles of body mass index (BMI), weight increase, the amount of visceral fat, physical inactivity, and a poor diet, have been proposed to explain how obesity leads to cancer. The most prominent candidate mechanisms that have been suggested to mediate the above obesity-related factors leading to cancer are: 1.) increased insulin
resistance and insulin-like growth factor-I (IGF-I), 2.) sex hormones (mainly estrogen) that are produced in excess amounts in fat tissue, 3.) adipokines (hormones in fat cells) that stimulate or inhibit cell growth, 4.) chronic, low level inflammation which can lead to cancer, 5.) weakened immune response affecting the NF-κB pathway, and 6.) oxidative stress.47,48,49,50

With respect to leukemia specifically, there is a lack of knowledge regarding these mechanisms, with the exception of IGF-I. Merchav et al. studied IGF-I and found it present in leukemic cells.51 IGF-I promotes cell proliferation and the propagation of bone marrow progenitor cells, and inhibits apoptosis- a plausible mechanism by which obese subjects with elevated IGF-I levels would be at an increased risk of developing leukemia. Regarding epidemiologic studies identifying obesity as a risk factor for CLL specifically, several have found that increased BMI at various ages in adulthood was associated with increased risk of CLL/SLL.52,53,54,55 These findings are consistent with a general study done on leukemia incidence in Europe which found high BMI to be associated with increased CML, AML, ALL, and CLL rates.56,29

Physical activity and diet are two important components of energy balance in the body, and both influence body weight and thus BMI. Both physical activity and diet are behavioral risk factors that can be potentially modified to prevent overweight and obesity. Currently, the NCI, under the Transdisciplinary Research on Energetics and Cancer (TREC) initiative is sponsoring studies to understand the relationship between physical activity and the risk of developing cancer.57 However, as of now, no studies
under this program have been published that specifically investigated leukemia incidence.

Independent of the TREC program, there are a limited number of studies to date that have investigated physical exercise as a risk factor for leukemia development. One recent meta-analysis by Jochem et al. compiled results from seven studies conducted in the U.S. and one from Europe. Overall, the outcome of their random effects meta-analysis did not show any association between physical activity and the development of leukemia.\textsuperscript{58} However, one of the studies in their meta-analysis, the largest one by far (prospective cohort of n=493,188 subjects), by Kabat et al. (2013) found a significantly decreased incidence of CML with higher level of physical activity, after adjusting for other covariates.\textsuperscript{59} For CLL, so far only one study has been able to show that increased physical activity could reduce incidence.\textsuperscript{60,61}

With respect to diet, there have only been a few studies to date that have shown association with leukemia. One large cohort that examined dietary exposures was the Iowa Women’s Health Study.\textsuperscript{25} Despite the large number of women studied (n=35,221), only a relatively small number developed leukemia (n=138), which could explain the weak associations observed with dietary factors. Nevertheless, their results did show a trend (P=0.08) for diets high in vegetables (all types of vegetables combined) being inversely associated with leukemia development. This suggests that by consuming moderate to high amounts of vegetables of any kind reduces a woman’s risk of leukemia by 50%. So far, this is the only study that points to the protective effect of vegetable consumption for leukemia development and several other studies have failed to show
similar results. The largest study to date focusing on dietary factors and CLL/SLL incidence was done by Tsai et al. in 2010. They did not show any association between these factors and CLL/SLL incidence, despite having a large number of cases—although the lack of longitudinal dietary data in this study could potentially limit their findings. With regards to the possible adverse impact of meat intake, Ma et al. investigated a range of meat consumption/preparation factors focusing on AML and found an association with higher meat intake and the risk for AML. However, their study could not identify a clear relationship between specific cooking methods or doneness of meat. For CLL/SLL specifically, meat intake did not seem to affect leukemia development in the Iowa Women’s Health Study, or in the Tsai et al. study.

Tea drinking, specifically green tea, is another dietary factor besides vegetable consumption that has recently been studied for its potential protective effect on the risk of leukemia. Zhang et al. conducted a hospital based case-control study in southeast China where regular green tea consumption is frequent with nearly half the population consuming green tea on a regular basis. Their findings indicate a significant dose-response relationship between increased green tea drinking and the reduction of risk in developing ALL, CML, and CLL/SLL, but not AML.

Overall, although increased BMI in adulthood appears to be a convincing risk factor for CLL/SLL, there remains a need to further evaluate the role of physical activity and dietary factors in the development of CLL/SLL in order to confirm or disprove the above findings.
2.4.1.4 Hair Dyes

Hair dyes and their possible carcinogenicity have been studied since the 1990s as an initiative by the International Agency for Research on Cancer (IARC), a branch of the World Health Organization. These investigations were prompted by the mutagenic and carcinogenic effect of \( p \)-phenylenediamine (PPD, a type of aromatic amine) and related nitro compounds observed in rats and mice. PPD has been and remains the main active ingredient in permanent and semi-permanent hair dyes. Exposure of PPD from hair dyes can occur through the skin of the scalp. From laboratory experiments, it appears that PPD in combination with hydrogen peroxide (a mixture commonly used as a method of hair coloring, especially for women with naturally darker hair color) is a potent carcinogen. This carcinogenic effect is likely to be amplified by long term, cumulative use of hair dyes.

Despite the aforementioned laboratory data, after their initial investigation, the IARC study found personal hair dye use to be only weakly associated with certain cancers (bladder, breast, skin, and lung), and found its association with hematopoietic neoplasms, including leukemia, inconclusive. As a consequence, several studies have since been conducted specifically focusing on the potential effects of hair dye use as a risk factor for various types of leukemia. Correa et al., published the largest literature review up to the present time on this topic, in 2000. They compared 11 epidemiologic studies that explored hair dye use as exposure for leukemia. Data collection for these studies spanned from 1976 until 1994 and mainly included women. To a variable extent,
the publications considered categories of: 1) permanent and semi-permanent hair dyes; 2) dark and light hair colors; 3) frequent and infrequent users; 4) long term users (10 years or more) and; 5) short term users (less than 10 years). However, the exposure assessment according to these categories was not standardized and many of them considered only one of the aforementioned factors or none at all. Despite the lack of consistency and methodological limitations across these studies, the authors found and confirmed evidence of at least a weak association of hair dye use increasing incidence of ALL, AML, and CLL- but not CML. In addition, Slager et al. focusing on CLL, found that both hairdresser occupation and higher frequency of personal hair dye use associated with incidence.

Generally, it appears that darker and more permanent hair dyes have a stronger association with leukemia incidence.\textsuperscript{69,70} In one of these studies, Grodstein et al. estimated only about a 10% prevalence of dark permanent hair dye use among women, which could explain why these studies have been severely underpowered, especially considering the outcome (i.e., leukemia incidence in the population) is rare as well.\textsuperscript{71} Considering these epidemiologic observations, albeit weak, in light of the known carcinogenic effects of PPD in combination with hydrogen peroxide, long-term hair dye use is a possible and modifiable risk factor for leukemia.

\textbf{2.4.1.5 Nonsteroidal Anti-inflammatory Drugs}

Although Nonsteroidal Anti-inflammatory Drug (NSAID) use is an active area of
epidemiologic studies, so far, only a couple of studies have been published exploring NSAID use on leukemia incidence.\textsuperscript{17,72,73} With respect to other cancers, specifically in solid tumors, more extensive research has been done. These point to certain NSAIDs leading to reduced cancer incidence. The strongest association has been found with colon cancer.\textsuperscript{74} Likewise, there is some evidence for certain NSAIDs being preventative in lung, breast, prostate, stomach, and esophageal cancers.\textsuperscript{75,76,77,78,79} The NSAIDs identified in these studies as exhibiting a protective effect were aspirin, ibuprofen, and COX-2 inhibitors- most likely due to their anti-inflammatory properties.

Published research studies to date aiming to evaluate the impact of NSAIDs use on leukemia incidence have consistently shown an increase in myeloid leukemia cases related to acetaminophen consumption.\textsuperscript{17,18,80} In addition Ross et al. found an important dose-response type of relationship between the amount of tablets taken per week and the incidence of AML and CML, in particular for females. They pointed to an over two-fold increase in the odds of leukemia for women taking seven or more tablets of acetaminophen per week on a regular basis, as opposed to almost no treatment effect at lower doses. Moreover, there are biological hypothesis, supported by both cell line and animal experiments that elucidate acetaminophen’s unfavorable activity in the mechanism of myeloid leukemia development. Acetaminophen is thought to increase chromosomal damage and inhibit ribonucleotide reductase, therefore also inhibiting DNA synthesis and repair.\textsuperscript{81,82,83,84} Additionally, animal studies have shown that at high doses (levels that result in hepatotoxicity), acetaminophen is genotoxic and causes bone marrow toxicity in rats.\textsuperscript{81} Relative to acetaminophen’s role in CLL/SLL specifically, there have
not been any sufficiently large studies published to date.

With respect to NSAID use, other than acetaminophen, for example aspirin, ibuprofen, naproxen, COX-2 inhibitors, the conclusions of these leukemia studies do not seem very consistent. Ross et al. showed that there appears to be a decreased risk for leukemia – specifically for myeloid types (AML and CML) - in women who use regular or extra strength aspirin. It is hypothesized that the anti-inflammatory properties of aspirin contribute to its protective effect for leukemia, particularly it has been found that aspirin can induce apoptosis in AML cell lines. In contrast, researchers did not observe any statistically significant associations for ibuprofen or Cox-2 inhibitors in either sex or in any of the leukemia subtypes. Due to the hypothesis of chronic inflammation’s role in the initiation of CLL/SLL, there is a need for studies investigating the potential protective role of NSAIDs for CLL/SLL, other than acetaminophen. One recent meta-analysis by Ye et al. points to the protective effect of aspirin use on CLL/SLL incidence, supporting the above hypothesis.

2.4.1.6 Hormone Therapy

Due to the wide-spread use of hormone therapy (HT) by post-menopausal women in recent decades, multiple studies have investigated the possible elevated risk of various cancers, potentially contributed by increased estrogen levels. To date, studies have indicated elevated risk for breast, ovarian, and endometrial cancers among women using HT. With respect to the risk of leukemia associated with HT use, studies have been
scarce and inconclusive so far. Since estrogen receptors are present on some hematopoietic cells, it is reasonable to suspect that there would be an association of HT use and leukemia incidence. Ross et al. investigated this potential relationship in the Iowa Women’s Health Cohort, but did not observe any effect of HT on leukemia incidence overall or by leukemia subsets. In addition to other leukemia types, their study only included 87 CLL cases, which is likely too few for this type of analysis. Therefore, additional research should further explore the effect of HT on CLL/SLL rates.

2.4.2 Environmental Risk Factors for Chronic Lymphocytic Leukemia/ Small Lymphocytic Lymphoma

A variety of potential occupational and environmental risk factors have been investigated and are the subject of ongoing research for their potential association with leukemia incidence. Thus far, ionizing radiation, benzene exposure, and treatment with cytostatic drugs have been documented for their causal association with leukemia in general, but not necessarily with CLL/SLL. In addition, certain viral infections have been implicated to lead to particular types of leukemia. Agricultural chemical use, particularly pesticides, has also been the subject of evaluations for its potential association with leukemia, with some recent research focusing on CLL/SLL.
2.4.2.1 Radiation

There are multiple forms of ionizing radiation – these can be categorized into two main types: man-made and natural. Man-made radiation can stem from medical radiation (such as during cancer therapy for both patient and operator), occupational radiation, nuclear power production, and testing and production of nuclear weapons. Natural radiation can be due to cosmic radiation or radiation from other natural sources, such as the soil (ex. radon). In addition, ultraviolet (UV) radiation – a non-ionizing radiation exposure – is of consideration in cancer epidemiology.

Ionizing radiation is a well-documented, adverse risk factor for leukemia. This is mainly true for AML, ALL, and CML. Chronic lymphocytic leukemia is considered a non-radiosensitive leukemia. Although not a true environmental exposure, but similar in effect to environmental radiation, patients receiving radiation therapy for cancer is one of the most frequent sources of ionizing radiation exposure and one that has been shown to result in AML, ALL, or CML, but not CLL. Leukemia developing from medical radiation is medically referred to as “therapy related leukemia.” This term is also used for leukemia resulting from previous chemotherapy exposure (see section 2.3.2.2. below). The amount of risk of leukemia following radiation therapy depends on several factors, including dose of radiation, the area of the body treated, and the age of the patient at time of the radiation treatment. Another mode of medical radiation leading to leukemia is one that affects the operator (ex. radiologist or nurse) of the radiation equipment – a type of occupational exposure. Studies exploring this type of exposure have not always been
conclusive. Other types of occupational radiation exposure involve workers at nuclear facilities, military facilities, or uranium mines. Of these, cleanup workers at nuclear facilities after a nuclear disaster had the strongest association with developing leukemia. A study by Romanenko et al. in 2008 found a linear dose-response relationship between increasing exposure to radiation and the risk of leukemia among cleanup workers, specifically ALL, following the Chernobyl nuclear accident of 1986. Studies conducted at nuclear power plants or at nuclear weapon facilities without incidents have typically shown weaker, but still often statistically significant, association of radiation and leukemia incidence (excluding CLL).

Natural sources of ionizing radiation are cosmic rays, gamma rays from naturally occurring radioisotopes such as potassium-40, radionuclides in food, and inhaled isotopes of radon. Studies evaluating these exposures for leukemia incidence have mainly been conducted with the aid of extrapolating radiation dose data from higher exposure settings. They do, however, suggest that even at lower doses, cumulative exposure over a period of many years elevates the risk of leukemia- particularly AML, ALL, and CML.

Ultraviolet radiation has been a subject of investigation for its potential positive association with leukemia incidence. However, to date, studies have yielded mixed conclusions regarding a positive association of UV exposure with leukemia. Moreover, several studies – primarily conducted in Northern Europe, the U.S., and Australia among Caucasian subjects- have found various measures of UV radiation to be protective against CLL development. There are a couple of possible mechanistic theories in
support of this finding. One is that UV radiation can modulate the immune system which alter T-cells (CD4+, CD8+, and natural killer cells) in a way that reduce the risk of CLL/SLL development. Another possible mechanism has to do with UV exposure inhibiting leukemogenesis through an increase in vitamin D production, and thereby reducing the risk of CLL/SLL- similarly to what has been observed for some solid tumors (ex. breast, prostate, colon, and ovarian cancers).

2.4.2.2 Secondhand Smoke

The evaluation of secondhand smoke as a risk factor for adult leukemia – unlike for solid tumor cancers- has not been studied very extensively to date. Chemicals and their carcinogenic properties in the air due to secondhand (also known as environmental or involuntary) smoke are essentially identical to direct exposure to firsthand smoking. There are two ways secondhand smoke can enter the exposed individual: through mainstream smoke and side stream smoke. Mainstream smoke gets directly inhaled into one’s lungs, while side stream smoke is the smoke at the end of the burning cigarette. About 15% of secondhand smoke can be attributed to mainstream smoke and 85% to side stream smoke- the latter which tends to contain more of certain carcinogens.

The most important study on secondhand smoking exposure and its relationship to leukemia was conducted in Canada by Kasim et al. They performed a large population based case-control study on lifetime non-smokers, using data from 1994-1997. The study assessed two sources of secondhand smoke exposure: residential and occupational. They
looked at all types of leukemia and although did not find an association of secondhand smoke with most types, they did with CLL. The association [odds ratio (OR) over two-fold] with CLL was very clear, with a significant dose-response relationship, for both residential and occupational exposure. They found the highest risk from occupational exposure- consistent with similar findings in lung cancer. Notably, this finding for CLL is not consistent with the modest protective effect observed for firsthand smoke – which emphasizes the need to evaluate secondhand smoking exposure as a separate factor from firsthand smoking. Other than this study, the remainder of the research on environmental smoke and leukemia has been focused on children and parental smoking exposure or on maternal smoking during pregnancy and the risk of childhood leukemia- with mixed conclusions.  

2.4.2.3 Pollution

The effect of environmental pollution on leukemia incidence has been studied world-wide in numerous industrialized countries. Generally, these studies have taken place in residential areas that are in the vicinity of certain industries that are known, or at the least suspected to be, associated with leukemia with respect to occupational exposure. It is not yet clear if pollution in the area would impact the health of the general community. Pollution of this type can affect the air, the water, the soil – and thereby the food supply, in these areas.

One study, evaluating the environmental pollution for populations living close to
an oil refinery emitting carcinogenic volatile organic compounds (VOCs, such as benzene) in Sweden, found a significant increase in leukemia incidence compared to an unaffected area. Although this study showed a significant excess of leukemia cases in the affected area, the authors were not able to completely attribute this finding to the amount of VOCs they measured. Hence, they could not establish a clear explanation for their findings.

A large cohort study recently conducted in China examined the effects of environmental pollution caused by pentachlorophenol (PCP) on the community. Pentachlorophenol is used as a pesticide, herbicide, and in the treatment of wood. Its production and resulting toxic waste can disseminate into the environment in various ways. This study assessed rates for various cancer types, including leukemia, among long-term residents in an area known to be polluted by PCP. The study found a significant correlation of PCP exposure and leukemia (all types analyzed together) incidence, in fact the association with leukemia was the strongest of all the cancers diagnosed.

Metal production and processing has been another area of investigation with respect to its polluting effects and leukemia. There have been multiple studies conducted in Spain, including one study specifically focusing on the metal industry’s effect on leukemia in neighboring areas. The authors were only able to assess leukemia mortality and not incidence - nevertheless found an excess in leukemia (all types combined) mortality among residents living close to metal production and processing industries.
A more general, case-control study on environmental pollution and its effects on leukemia (all types combined) risk was done recently in an industrial region of northern Italy. The area studied contained pollutants from a mixture of sources: a fossil fuel power plant, a coke oven, and a couple of chemical plants. The authors observed an elevated risk for leukemia in the exposed areas but the results did not reach statistical significance, most likely due to small sample size.\(^{107}\)

The aforementioned studies indicate that pollution from industries known to produce carcinogens generally contribute to increasing leukemia risk for populations living in the vicinity. However, due to the difficulty of measuring exposure, causality may be difficult to assess. In addition, these studies did not differentiate their exposure assessment by leukemia subtype, in particular there are no specific data relative to CLL/SLL risk.

### 2.4.2.4 Pesticides and Herbicides as Occupational Exposures

Pesticide and herbicide exposure are potential occupational risk factors for cancers, including leukemia in general. These exposures originate from both agricultural applications (farming industry and private residences) as well as the manufacturing of these chemicals. Despite the heterogeneity in products, for a large proportion of chemicals used as pesticide and herbicides, there is biological evidence of carcinogenicity. There have been several investigations published over the past couple of decades on this subject, evaluating various chemicals either separately or in combination.
However, these epidemiologic studies have not typically been powered adequately to detect significance in possible associations and there is a lack of consistency in the characterization of exposure. Moreover, most of the research on this subject did not distinguish between leukemia subtypes. Below are some examples of studies that looked at a large number of pesticide/herbicide exposed subjects. These studies mainly considered people working as applicators in the farming sector. However, some recent studies investigated people who do not work in agricultural jobs, nevertheless have potentially increased pesticide exposure by living on a farm.

Organophosphate pesticides (also called fonofos) and their impact of health have been an area of investigation. The largest study to examine associations of fonofos exposure with leukemia was conducted by Mahajan et al.\textsuperscript{108} As part of the Agricultural Health Study, they studied a large cohort of pesticide applicators (n=45,372) from Iowa and North Carolina and collected cancer incidence data on them. Of all the cancers evaluated, only leukemia turned out to be associated with fonofos exposure in this cohort. For applicators in the highest category of lifetime exposure to this pesticide, the relative risk of being diagnosed with leukemia was about two and a half times higher, compared to the lowest lifetime exposure category, with a significant linear trend in the risk of leukemia increasing with higher exposure levels. Although the study was not adequately powered to examine leukemia by subtypes, the authors did find the same trends of increased incidence in AML, CML, and CLL (ALL could not be evaluated).

Similar to the above study, Alachlor (trade name: Lasso), an herbicide used in the
production of soy, corn, and peanuts has also been evaluated for its association with cancer in multiple studies. This herbicide has been widely used since 1969, despite its 1985 categorization by the U.S. Environmental Protection Agency as a probable human carcinogen. As part of another study under the Agricultural Health Study, Lee et al. assessed cancer incidence among Alachlor applicators. They found a significant increase in the risk of leukemia (all types together), as well as multiple myeloma-associated with high lifetime exposure of Alachlor. No other significant associations with other cancers were found.

Another recent study evaluating pesticide use and cancer incidence was done in Turkey. Uysal et al. conducted their study specifically in an area of Turkey with large scale greenhouse farming—a type of agricultural practice requiring intensive use of pesticides. The researchers found a significant association between increased incidence of leukemia (ALL specifically) among the population living in the area and pesticide use. In addition, the authors also saw associations with multiple myeloma and malignant melanoma. The study did not differentiate among type of pesticides used.

In their large, pooled analysis Slager et al. confirmed a positive relationship of working on a farm and CLL/SLL incidence. In addition, they demonstrated a positive (although not statistically significant) association between living on (or near) a farm and the risk of CLL. Further studies are needed evaluate this novel possible relationship, preferably using more precise measures of exposure. Furthermore, residential home pesticide application could also confer an increased risk for CLL/SLL and would be of
interest to investigate.

Overall, the inconsistencies in association between pesticide exposure and leukemia are most likely due to the large heterogeneity of the chemicals investigated, as well as the characterization of exposure measurements across the studies. In addition, the small number of cases by leukemia subtype is a limiting factor in assessing which subtypes are related to which pesticide exposures.

2.4.2.5 Benzene as Occupational Exposure

Benzene is a widely recognized leukemogen. In addition to being a major component of cigarette smoke (2.3.1.1 and 2.3.2.3), workers in certain industries chronically exposed to benzene are at elevated risk of developing MDS (which often leads to AML) or AML directly. Industries with high levels of benzene exposure are tire manufacturing, oil refineries and other gasoline related industries, chemical plants, and shoe factories. Originally it was thought that the characteristics of AML resulting from benzene exposure closely resembled therapy-related AML, however more recent research shows that in fact they more look like de novo AML, based on detailed cytogenetic, hematologic, and epidemiologic investigation.\textsuperscript{111,112}

The Pliofilm rubber worker cohort composed of workers employed by the Goodyear tire factories in Ohio from the period of 1936 to 1975, has been the key to
raising awareness and to establishing health-based standards for benzene by the U.S. EPA and the Occupational Safety and Health Administration (OSHA).\textsuperscript{113,114,115} Paxton et al. in 1993 assessed the mortality rate for leukemia due to benzene exposure in this cohort, in which all leukemia (all types combined) deaths occurred for workers exposed to benzene prior to 1950, during a period when there were less measures of industrial hygiene and greater levels of exposure. They found a statistically significant standardized mortality ratio of 3.36 for workers exposed to benzene, with a clear dose response relationship between person-years of exposure and mortality from leukemia (all types combined). Workers in the highest category of exposure years had an SMR greater than ten. These data were subsequently updated and confirmed after additional follow-up. Overall, there appears to be a 50 parts per million (ppm) - years threshold of cumulative benzene exposure required for leukemogenesis.

The petroleum industry has been another site of research with respect to benzene exposure and leukemia. Multiple studies have been conducted investigating this exposure, based on petroleum worker cohorts from Canada, United Kingdom, and Australia.\textsuperscript{116,117,118} A study recently published presents pooled analysis of data collected in these three countries.\textsuperscript{119} The authors found a relationship with benzene exposure and AML, although this was weaker than observed in some earlier studies, which had higher benzene exposures. In this pooled study however, the authors did identify a subgroup of petroleum workers, the ones driving the tanks at terminals, who had an over two-fold odds of developing AML compared to regular petroleum workers- which observation confirms a previously found dose-response relationship.
Hayes et al. in 1997 published a large case-control study conducted in China. They compared hematological cancer rates between workers in various industries (painting, printing, chemical, shoe) exposed to benzene (n=74,828) compared those in similar industries but not exposed (n=35,805).\textsuperscript{120} They found a dose-response relationship between the level –quantified as cumulative ppm- of benzene exposure and the risk of AML and MDS, demonstrating a significant elevation in the risk at lower cumulative doses as well.

Overall, workers at wide range of exposures to benzene through various occupations have been found to have an elevated risk for AML and MDS, confirming the leukemogenic potential of this chemical. Benzene exposure does not appear to be a risk factor for other types of leukemia, for example CLL/SLL. However, since benzene is found in cigarettes, it may have an indirect effect on CLL/SLL, particularly in secondhand smoke that appears to be a risk factor for CLL/SLL as mentioned above.

2.4.2.6 Viruses

Viral infection as a risk factor for leukemia can be considered both as an environmental and as a biological risk factor. In humans, so far the only virus that has been confirmed to cause leukemia is the Human T-cell lymphotropic virus type 1 (HTLV-I), also known as adult T-cell lymphoma virus type 1. HTLV-I is a virus that is associated with adult T-cell leukemia/lymphoma (ATL)- a highly aggressive malignancy. The causative relationship between HTLV-I and ATL has been well established, as the
HTLV-I proviral genome is detectable in blasts of ATL patients.

Up to five percent of the people infected with the virus are thought to develop ATL. The main route of transmission of HTLV-I is mother to child via breastfeeding. Other, less common means of transmission are by blood contact (transfusions, sharing of needles/syringes) and sexual intercourse. There are important geographical differences in the prevalence of HTLV-I infection worldwide, with Japan having the highest rates of infection (up to 37% of the population infected in some localities), followed by sub-Saharan Africa, the Caribbean islands, Colombia, and Brazil. The onset of ATL on average occurs in the second or third decade of life, however the mean age of onset varies globally: fourth decade in Brazil and in the Caribbean; fifth decade in Japan.\textsuperscript{121}

ATL is a heterogeneous disease that can be grouped into four subtypes: lymphoma, acute, chronic, and smoldering type. Each of these subtypes has different diagnostic criteria and clinical outcome. Diagnostic criteria are a function of lymphocyte count, percent of atypical lymphocytes, Lactate dehydrogenase (LDH), and calcium levels, in addition to evaluation of extramedullary involvement. Patients with the acute leukemia and lymphoma type have the worse overall survival, followed by the chronic type and then the smoldering type- but all subtypes have poor prognosis relative to other types of leukemia.\textsuperscript{122}

Another virus suspected to be associated with leukemia, CLL specifically, is Hepatitis C virus (HCV).\textsuperscript{123} The presence of HCV in CLL patients is typically correlated
with a more advanced disease. It is hypothesized that HCV infection leads to deregulation of the immune system among B cells. According, to recent estimates, HCV infection increases the odds of CLL/SLL incidence by about two-fold.\textsuperscript{35,124}

2.4.3 Biological Risk Factors for Chronic Lymphocytic Leukemia/ Small Lymphocytic Lymphoma

A family history of hematological cancers appears to be a frequent biological association relative to CLL/SLL development. Adult height, an anthropologic measurement and a surrogate of several biological systems (ex. genetic, nutritional, immunological, and hormonal), was recently shown to be associated with CLL/SLL incidence, with tall subjects (in the fourth quartile) having an increased risk of CLL/SLL.

2.4.3.1 Familiar Link and Genetic Diseases

While for most leukemia cases, no familial link has been indicated or suspected, associations have been noted with a small subset of cases. In addition, certain genetic abnormalities – mainly present at birth – have been found to have a strong relationship with the development of various types of leukemia.

With respect to familial link, it has been established that having first-degree relatives with a hematologic disease elevates the risk of developing leukemia, mainly CLL/SLL.\textsuperscript{125} Furthermore, having an identical twin with a history of AML or ALL is associated with an increased risk for AML or ALL, with about 20% of the other twins
developing these acute types of leukemia. \textsuperscript{126,127} Additionally, a history of breast cancer in sisters (but not in mothers) is associated with both AML and ALL. This may be due to specific germ-line mutations that can run in families that could cause both breast cancer and leukemia clusters. Certain other exposures, such as smoking, ionizing radiation, aromatic hydrocarbon exposure, added to a family history of breast cancer can further increase the odds of developing leukemia (studies so far only done on AML and ALL). \textsuperscript{13, 128}

Regarding genetic diseases, certain specific ones are thought to play a role in the development of leukemia. The majority of these aberrations such as Down syndrome, Shwachman syndrome, Li-Fraumeni syndrome, and Neurofibromatosis are primarily involved in pediatric leukemia. Yet, a few genetic abnormalities such as Fanconi anemia, Ataxia telangiectasia, and Kostmann syndrome are also related to leukemia (various types, including CLL) development in adulthood, in addition to their association with pediatric leukemia. These genetic abnormalities often have milder phenotypes and can be successfully treated, resulting in longer life expectancy. \textsuperscript{129,130,131,132,133,134,135}

A rare but serious condition, Fanconi anemia is an X-linked autosomal recessive disorder that results in progressive pancytopenia, and, along with a wide range of clinical symptoms, it is associated with chromosomal instability and increased risk for cancer, namely AML. \textsuperscript{136,137} It affects 1 per 350,000 births and about half the patients with this disorder will develop AML by the time they reach 40 years of age. Patients typically require allogeneic stem cell transplantation.
Ataxia telangiectasia is an autosomal recessive immunodeficiency disorder that affects different organs. It is less frequent than Fanconi anemia, affecting 1 per 400,000 births (although some estimates are even lower) and of the approximately 20% of patients with this disease who develop cancer, 10% develop lymphoid malignancies (ALL or lymphoma) specifically. Rarely, other types of leukemia have also been associated with Ataxia telangiectasia.¹³⁸

Kostmann syndrome, a form of congenital neutropenia, is another rare autosomal recessive disorder that is associated with leukemia (AML and MDS). It affects about 1 in 1,000,000 births.¹³⁹ It is caused by a defect of the granulocyte colony-stimulating factor receptor (GCSFR) gene on chromosome 1p35-p34.3. Patients can be treated with granulocyte colony-stimulating factor (G-CSF) to ease and delay symptoms. The incidence of AML or MDS in Kostmann syndrome after ten years of G-CSF treatment is about 20%.¹³³,¹⁴⁰ Kostmann syndrome does not seem to be associated with CLL.

2.4.3.2 Height

Height – through different pathways- is a measure that may associated with CLL, in addition to the other most notable anthropometric measure, weight. The largest pooled study to date done by Slager et al. has observed a positive relationship between adult height and an increase in CLL incidence.³⁵ Engel and et al. also found a similar relationship with height and CLL.⁵⁶ While there is no clear explanation for this association, there are some proposed mechanisms.
Height is a function of genetics, nutrition (early in life), immune system, and hormone levels (including growth hormones). Relative to the immunology theory, short stature is thought to be affected by frequent childhood infections - which in turn result in a stronger immune system in adulthood: a potentially protective factor against leukemogenesis. On the contrary, taller stature is associated with an increased availability of nutrition during childhood, enabling elevated growth hormones exposure (for example insulin-like growth factor-1 (IGF-1)). IGF-1 has been shown to both stimulate B-cell proliferation and to inhibit apoptosis. As all of these factors potentially play a role in leukemia development, height may be a well-suited surrogate of many biological systems in leukemogenesis. Additional studies are needed to confirm this finding.

2.4.4 Summary of Risk Factors for Chronic Lymphocytic Leukemia/ Small Lymphocytic Lymphoma

Overall, there have been several risk factors investigated for leukemia and NHL specifically, however only a few studies have focused on CLL/SLL incidence relative to these risk factors. Table 1.1 below summarizes the existing research on leukemia risk factors and indicates if any of the findings pertain to CLL/SLL specifically. Only a handful of investigated potentially modifiable risk factors showed association with CLL/SLL. These are green tea consumption, hair dye use, radiation, secondhand smoke, and certain viral infections. The associations with these risk factors were relatively weak and typically inconsistent across studies. In addition, none of these findings explain the lower risk observed among women or why CLL/SLL is more common in industrialized...
countries. Therefore, we sought to conduct a more comprehensive investigation of modifiable risk factors focused on CLL/SLL incidence in postmenopausal women. Specifically, in our study we had well characterized and adequate data on a wide range of parameters related to personal behaviors (e.g. diet, drinking, and exercise habits), hormone therapy, and pesticide exposures, and hence we were able to analyze these potential risk factors for CLL/SLL. Information on past NSAID use and hair dye use was not sufficiently available for our study, thus we were not able to investigate these factors.
Table 1.1 below summarizes research on various exposures for CLL/SLL

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>International Agency for Research on Cancer</td>
<td>2002</td>
<td>Literature review</td>
<td>Cigarette smoking</td>
<td>Smoking is implicated as an important and probable causal risk factor for leukemia, but not for CLL.</td>
<td>No</td>
</tr>
<tr>
<td>Sandler et al. [36]</td>
<td>1993</td>
<td>Case-control</td>
<td>Cigarette smoking</td>
<td>Subjects 60 years or older who had any smoking history, had a significant increase of acute leukemia incidence (2-fold in AML, 3-fold in ALL). Certain morphologic subtypes (French-American-British (FAB) classification M2 in AML, FAB classification L2 in ALL) and specific cytogenetic abnormalities had increased associations with smoking.</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2) Exposure to alcohol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rauscher et al. [27]</td>
<td>2004</td>
<td>Case-control</td>
<td>Alcohol Consumption</td>
<td>Showed an inverse association of acute leukemia with light to moderate beer intake, as well as a positive association of acute leukemia with moderate to heavy wine intake.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Jianguang et al. [43]</td>
<td>2014</td>
<td>Cohort</td>
<td>Alcohol Consumption</td>
<td>Subjects with alcohol use disorders had a reduced incidence of leukemia, supporting a possible protective mechanism. However, it was conducted in a special population of subjects with heavy alcohol consumption, therefore, these results may not be applicable to the general population.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Mirjam et al. [44]</td>
<td>2014</td>
<td>Cohort</td>
<td>Alcohol Consumption</td>
<td>Suggests a possible positive relationship between alcohol (any type) consumption and the incidence of leukemia.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Diaz et al. [45]</td>
<td>2002</td>
<td>Literature review</td>
<td>Alcohol Consumption</td>
<td>Provides support for moderate alcohol consumption to improve immune response but for heavy alcohol use to impair immune function.</td>
<td>Evidence for leukemia (general)</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>3) Obesity/Physical Activity/ Diet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merchav et al. [51]</td>
<td>1998</td>
<td>Laboratory</td>
<td>Obesity</td>
<td>IGF-I is present in leukemic cells.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Kabat et al. [59]</td>
<td>2013</td>
<td>Cohort</td>
<td>Lifestyle and Diet</td>
<td>An increased BMI was associated with an increased risk of CML. CML incidence decreased significantly with higher level of physical activity, after adjusting for other covariates. Found no association between diets high in vegetables and leukemia development.</td>
<td>No</td>
</tr>
<tr>
<td>Engeland et al. [56]</td>
<td>2006</td>
<td>Cohort</td>
<td>Lifestyle and Diet</td>
<td>High BMI was associated with increased CML, AML, ALL, and CLL/SLL incidence rates.</td>
<td>Evidence for all leukemia types</td>
</tr>
<tr>
<td>Jochem et al. [58]</td>
<td>2014</td>
<td>Meta-analysis</td>
<td>Lifestyle and Diet</td>
<td>Overall, no association between physical activity and the development of leukemia was found.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Ross et al. [25]</td>
<td>2002</td>
<td>Cohort</td>
<td>Diet</td>
<td>Diets high in vegetables (all vegetables) are inversely associated with leukemia development. Consuming moderate to high amounts of vegetables of any kind reduces a woman’s risk of leukemia by 50%. Meat intake did not seem to affect leukemia development.</td>
<td>Evidence for leukemia (general)</td>
</tr>
</tbody>
</table>
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Auth Year</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ma et al. [62]</td>
<td>2010</td>
<td>Cohort</td>
<td>Diet</td>
<td>No association between diets high in vegetables and leukemia development. Found an association with higher meat intake and the risk for AML.</td>
<td>No</td>
</tr>
<tr>
<td>Zhang et al. [64]</td>
<td>2008</td>
<td>Case-control</td>
<td>Diet</td>
<td>Findings indicate a significant dose-response relationship between increasing green tea drinking and the reduction of risk in developing ALL, CML, and CLL – but not AML.</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>4) Exposure to Hair Dyes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correa et al. [68]</td>
<td>2000</td>
<td>Literature review</td>
<td>Hair Dye Exposure</td>
<td>Evidence of a weak association of hair dye use increasing incidence of ALL, AML, and CLL- but not CML.</td>
<td>Yes</td>
</tr>
<tr>
<td>Miligi et al. [69]</td>
<td>2005</td>
<td>Case-control</td>
<td>Hair Dye Exposure</td>
<td>Darker and more permanent hair dyes have a stronger association with ALL and CLL/SLL incidence.</td>
<td>Yes</td>
</tr>
<tr>
<td>Rauscher et al. [19]</td>
<td>2004</td>
<td>Case-control</td>
<td>Hair Dye Exposure</td>
<td>Darker and more permanent hair dyes have a stronger association with AML and ALL incidence.</td>
<td>No</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5) Exposure to Nonsteroidal Anti-inflammatory Drugs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross et al. [17]</td>
<td>2011</td>
<td>Case-control</td>
<td>Nonsteroidal Anti-inflammatory Drugs</td>
<td>Showed an increase in AML and CML incidence related to acetaminophen consumption. Women taking 7 or more tablets of acetaminophen per week on a regular basis had an over 2-fold increase in leukemia incidence, as opposed to almost no effect for lower doses. Found women who use regular or extra strength aspirin have a decreased risk for AML and CML. Found no statistically significant associations for ibuprofen or Cox-2 inhibitors in either sex or in any of the leukemia subtypes.</td>
<td>No</td>
</tr>
<tr>
<td>Walter et al. [18]</td>
<td>2011</td>
<td>Cohort</td>
<td>Nonsteroidal Anti-inflammatory Drugs</td>
<td>Increase in AML and CML incidence related to acetaminophen consumption. No associations with other NSAIDs were found.</td>
<td>No</td>
</tr>
<tr>
<td>Robak et al. [80]</td>
<td>2008</td>
<td>Literature review</td>
<td>Nonsteroidal Anti-inflammatory Drugs</td>
<td>Increase in acute leukemia incidence related to acetaminophen consumption. Inverse correlation with acute leukemia and aspirin use.</td>
<td>Evidence for leukemia (general)</td>
</tr>
</tbody>
</table>
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>6) Hormone therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross et al. [88]</td>
<td>2005</td>
<td>Cohort</td>
<td>Hormone Therapy</td>
<td>No evidence of Hormone therapy on the incidence of AML or CLL- but only 87 cases of CLL were studied.</td>
<td>No</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Cancer Society [90]</td>
<td>2012</td>
<td>Literature review</td>
<td>Radiation</td>
<td>Patients receiving radiation therapy for cancer is one of the most frequent sources of ionizing radiation exposure and one which has been shown to result in leukemia.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Polychronakis et al. [15]</td>
<td>2013</td>
<td>Literature review</td>
<td>Radiation</td>
<td>Medical radiation affecting the operator of radiation equipment is associated with leukemia. Of other types of occupational radiation exposure, cleanup workers at nuclear facilities after a nuclear disaster had the strongest association with developing leukemia.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Romanenko et al [91]</td>
<td>2008</td>
<td>Case-control</td>
<td>Radiation</td>
<td>Linear dose-response relationship between increasing exposure to radiation and the risk of leukemia among cleanup workers, specifically ALL and CLL, following the Chernobyl nuclear accident of 1986.</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Romanenko et al [92]</td>
<td>2008</td>
<td>Case-control</td>
<td>Radiation</td>
<td>Linear dose-response relationship between increasing exposure to radiation and the risk of leukemia among cleanup workers, specifically ALL and CLL, following the Chernobyl nuclear accident of 1986.</td>
<td>Yes</td>
</tr>
<tr>
<td>Metz-Flamant [93]</td>
<td>2013</td>
<td>Cohort</td>
<td>Radiation</td>
<td>Weak association of radiation and leukemia mortality (all types except CLL) in subjects at nuclear facilities without incident.</td>
<td>No</td>
</tr>
<tr>
<td>Richardson [94,95]</td>
<td>2007</td>
<td>Cohort</td>
<td>Radiation</td>
<td>Mortality from leukemia was significantly elevated among workers at the Savannah River Site exposed to radiation.</td>
<td>Evidence for leukemia (general)</td>
</tr>
</tbody>
</table>
Table 1. Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilbert [96]</td>
<td>2009</td>
<td>Pooled analysis</td>
<td>Radiation</td>
<td>Showed that even at lower doses, given cumulative exposure over a period of many years to natural or artificial sources of radiation, the risk of leukemia is elevated.</td>
<td>Evidence for leukemia (general)</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>8) Secondhand Smoke</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kasim et al. [103]</td>
<td>2005</td>
<td>Case-control</td>
<td>Secondhand Smoke</td>
<td>Found a clear association between secondhand smoke and CLL, with a significant dose-response relationship, for both residential and occupational exposure. Found that the highest risk of CLL was from occupational exposure.</td>
<td>Yes</td>
</tr>
<tr>
<td>International Agency for Research on Cancer [34]</td>
<td>2002</td>
<td>Literature review</td>
<td>Secondhand Smoke</td>
<td>Mixed results concerning an association between parental smoking exposure or maternal smoking during pregnancy and the risk of childhood leukemia.</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>9) Pollution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barregard et al. [104]</td>
<td>2009</td>
<td>Cohort</td>
<td>Pollution</td>
<td>Significant increase in leukemia incidence for populations living close to an oil refinery – emitting carcinogenic volatile organic compounds - compared to an unaffected area. No analyses by type of leukemia.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Zheng et al. [105]</td>
<td>2013</td>
<td>Cohort</td>
<td>Pollution</td>
<td>Significant correlation between long term environmental Pentachlorophenol exposure and leukemia incidence.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Garcia-Pérez J et al. [106]</td>
<td>2010</td>
<td>Cohort</td>
<td>Pollution</td>
<td>Found an excess in leukemia mortality among residents living close to metal production and processing industries.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Parodi et al. [107]</td>
<td>2013</td>
<td>Case-control</td>
<td>Pollution</td>
<td>Elevated risk for leukemia in areas exposed to a fossil fuel power plant, a coke oven, and a couple of chemical plants. The results did not reach statistical significance, most likely due to small sample size.</td>
<td>Evidence for leukemia (general)</td>
</tr>
</tbody>
</table>
Table 1. Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>10) Pesticides and Herbicides as occupational exposures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mahajan et al. [108]</td>
<td>2006</td>
<td>Cohort</td>
<td>Pesticides</td>
<td>Leukemia incidence associated with fonofos exposure in a large cohort of pesticide applicators.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Lee et al. [14]</td>
<td>2004</td>
<td>Cohort</td>
<td>Herbicides</td>
<td>Significant increase in the risk of leukemia (all types together) - as well as multiple myeloma- associated with high lifetime exposure of Alachlor.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td>Uysal et al. [110]</td>
<td>2013</td>
<td>Cohort</td>
<td>Pesticides</td>
<td>Found a significant association between increased incidence of ALL among the population living in an area with large scale pesticide use. In addition, the authors saw associations with multiple myeloma and malignant melanoma.</td>
<td>No</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>11) Benzene as occupational exposure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paxton et al. [113,114,115]</td>
<td>1994</td>
<td>Cohort</td>
<td>Benzene</td>
<td>Statistically significant standardized mortality ratio of 3.36 for workers exposed to benzene, with a clear dose response relationship between person-years of exposure and mortality from leukemia (type not specified). Workers in the highest category of exposure years had an SMR greater than 10. Found a 50 parts per million (ppm) - years threshold of benzene exposure required for leukemogenesis.</td>
<td>Evidence for leukemia (general)</td>
</tr>
<tr>
<td></td>
<td>1996</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rushton et al. [119]</td>
<td>2014</td>
<td>Cohort</td>
<td>Benzene</td>
<td>Found a relationship with benzene exposure and AML. The subgroup of petroleum workers who drove the tanks at terminals had an over two-fold odds of developing AML compared to regular petroleum workers.</td>
<td>No</td>
</tr>
<tr>
<td>Hayes et al. [120]</td>
<td>1997</td>
<td>Case-control</td>
<td>Benzene</td>
<td>Dose-response relationship between the level – quantified as cumulative ppm- of benzene exposure and the risk of AML and MDS, demonstrating a significant elevation in the risk at lower cumulative doses.</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>12) Viruses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shimoyama et al. [122]</td>
<td>1991</td>
<td>Cohort</td>
<td>Viruses</td>
<td>Patients with the acute and lymphoma subtypes of ATL have the worst overall survival, followed by the chronic type and then the smoldering type- but all subtypes have poor prognosis relative to other types of leukemia.</td>
<td>No</td>
</tr>
<tr>
<td>De Sanjose et al. [124]</td>
<td>2008</td>
<td>Case-control</td>
<td>Hepatitis C</td>
<td>Hepatitis C viral infection increases the odds of CLL incidence by about two-fold.</td>
<td>Yes</td>
</tr>
<tr>
<td>Study area / Authors</td>
<td>Year</td>
<td>Type of study</td>
<td>Exposure</td>
<td>Main findings</td>
<td>Evidence for CLL/SLL?</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------</td>
<td>---------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>13) Familiar Link and Genetic Diseases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuttner [125]</td>
<td>1992</td>
<td>Cohort</td>
<td>Biological</td>
<td>First degree relatives of CLL patients have an elevated risk of developing some form of hematologic disease, mainly leukemia.</td>
<td>Yes</td>
</tr>
<tr>
<td>Rauscher et al. [126]</td>
<td>2002</td>
<td>Case-control</td>
<td>Biological</td>
<td>Found that having an identical twin with a history of AML or ALL is associated with an increased risk for AML or ALL, with about 20% of the other twins developing leukemia. A history of breast cancer in sisters (but not in mothers) is associated with both, AML and ALL.</td>
<td>No</td>
</tr>
<tr>
<td>Cortes et al. [127]</td>
<td>1996</td>
<td>Literature review</td>
<td>Biological</td>
<td>Having an identical twin with a history of AML or ALL is associated with an increased risk for AML or ALL.</td>
<td>No</td>
</tr>
<tr>
<td>Rauscher et al. [13]</td>
<td>2003</td>
<td>Cohort</td>
<td>Biological</td>
<td>A history of breast cancer in sisters (but not in mothers) is associated with both, AML and ALL. Certain other exposures, such as smoking, ionizing radiation, aromatic hydrocarbon exposure added to a family history of breast cancer can further increase the odds of developing leukemia.</td>
<td>No</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auerbach et al. [136]</td>
<td>1991</td>
<td>Literature review</td>
<td>Biological</td>
<td>Fanconi anemia is associated with an increased risk for AML. About half the patients with this disorder will develop AML by the time they reach 40 years of age.</td>
<td>No</td>
</tr>
<tr>
<td>National Cancer Institute. Bethesda, MD [138]</td>
<td>2014</td>
<td>Literature review</td>
<td>Biological</td>
<td>Approximately 20% of patients with Ataxia telangiectasia develop cancer, 10% develop lymphoid malignancies (ALL or lymphoma) specifically.</td>
<td>No</td>
</tr>
<tr>
<td>Germeshausen et al. [133]</td>
<td>2009</td>
<td>Cohort</td>
<td>Biological</td>
<td>The incidence of AML or MDS in Kostmann syndrome after 10 years of G-CSF treatment is about 20%</td>
<td>No</td>
</tr>
</tbody>
</table>

(continued)
Table 1.1 Reviewed articles on risk factors for CLL/SLL (continued)

<table>
<thead>
<tr>
<th>Study area / Authors</th>
<th>Year</th>
<th>Type of study</th>
<th>Exposure</th>
<th>Main findings</th>
<th>Evidence for CLL/SLL?</th>
</tr>
</thead>
<tbody>
<tr>
<td>14) Height</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slager et al. [35]</td>
<td>2014</td>
<td>Case-control</td>
<td>Height</td>
<td>Positive relationship between adult height and an increase in CLL incidence.</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(pooled analysis)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engeland et al. [56]</td>
<td>2006</td>
<td>Case-control</td>
<td>Height</td>
<td>Positive relationship between adult height and an increase in CLL incidence.</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 3: Methods

3.1 Overview

Our study is a secondary analysis using data from the WHI. The WHI is composed of two main elements: the Clinical Trials (CT) and the Observational Study (OS). Subjects enrolled in the WHI were 161,808 postmenopausal women ages 50-79 years, between 1993-1998, across 40 centers within the United States. The CT portion includes three randomized components: 1) the Hormone Therapy Trial (HT), 2) the Dietary Modification Trial (DM), and the 3) Calcium and Vitamin D Trial (CaD) [Figure 1.1]. Women eligible for the CT part could be randomized into either one, two, or all three of the trials, resulting in seven possible combinations of trial participation (i.e., HT alone, DM alone, CaD alone, HT+DM, HT+CaD, DM+CaD, HT+DM+CaD). Women who either did not meet eligibility criteria for the CT or were not willing to participate in it were invited to participate in the OS. Women in both the CT and the OS were to be followed for an average of 9 years.

The CT component has a partial factorial design that allowed for the evaluation of three hypotheses in a controlled setting of three randomized trials. The first arm (HT component hormone therapy [Estrogen plus Progestin arm and an Estrogen-alone arm]) was primarily hypothesized to reduce the risk of coronary heart disease and other
cardiovascular diseases, and secondarily to reduce the risk of hip and other fractures, with increased breast cancer risk as a possible adverse outcome. The second arm (DM component), which consisted of a low-fat diet, was hypothesized to prevent breast cancer and colorectal cancer and, secondarily, coronary heart disease. The third arm (CaD), which consisted of calcium and vitamin D supplementation, was hypothesized to prevent hip fractures and, secondarily, other fractures and colorectal cancer.

The Observational Study allowed for the evaluation of new risk factors, and their possible relationship to a variety of diseases and mortality. Baseline data collected for the OS was identical to those collected for the CT and included an extensive list of factors through physical measurements, interviews, and blood specimen collection/banking. Primary clinical end points of interest in the OS were breast and colorectal cancer, coronary heart disease and stroke, osteoporotic fractures, diabetes, and total mortality. Secondary endpoints comprise a large array of other diseases, including many cancers such as leukemia. Women in the OS were mailed annual questionnaires to update their disease and exposure status. In addition, about three years into the study, women returned to the clinic for blood work and physical evaluation. Of note, a high proportion (about 50%) of women participating in the OS had a history of HT use, as well as ongoing use and often declined to participate in the CT as they were already taking HT. Past and current/ongoing HT use in the OS cohort was documented in the WHI database.

Overall, the results of the CT revealed some important findings. Most notably, the HT trial’s “estrogen plus progestin” intervention arm pointed to some adverse health risks
(which resulted in early termination of this hormone therapy treatment arm), despite finding certain benefits. Particularly, the risk of stroke, pulmonary embolism, invasive breast cancer, and dementia all increased among women on this treatment. Benefits included reduced hip fractures and diabetes. Although the adverse effects were not as severe for the “estrogen only” arm in this study, they still resulted in an increased number of thromboembolic events for women assigned to this intervention. Women on these two hormone therapy arms have been followed after the interventions stopped and the general conclusion remains that the risk-benefit pattern due to hormone therapy does not support the use of these treatments in chronic disease prevention, and it should be used only for symptom management on an individual basis.\textsuperscript{148, 149} Importantly, these findings resulted in a substantial reduction of hormone therapy use among postmenopausal women in the U.S., which in turn also resulted in significant health care savings and reduction in breast cancer incidence.\textsuperscript{150}

The low fat dietary intervention arm of the DM trial ended up being less controversial. This dietary intervention showed a decreased risk for ovarian cancer.\textsuperscript{151} In addition, a non-significant trend was found for decreased invasive breast cancer incidence.\textsuperscript{152} However, the low fat dietary intervention did not show an effect in preventing endometrial or colorectal cancers, or cardiovascular disease.\textsuperscript{151, 153, 154}
The results from CaD trial did not support any of the study hypothesis: calcium and vitamin D supplementation was not associated with a decrease in hip fractures or colorectal cancer incidence. Nevertheless, exploratory data analyses showed lower vertebral fracture rates and lower incidence of in situ breast cancer incidence among women on the intervention arm.\textsuperscript{155}

**Figure 1.1 Venn Diagram of the WHI study design**

![Venn Diagram](image_url)

Note: Total Clinical Trials participants: N=68,132. Total WHI participants: N=161,808
3.1.1 Extension Studies

To allow for evaluation of longer term effects of interventions and to assess rare diseases, the follow-up period in the WHI was maximized by the implementation of two sequential extension studies (ES). At the end of the original follow-up period in 2005, women were asked to join the first WHI ES, for an additional 5 years of follow-up. In 2010, the current participants were invited to continue for an additional 5 years for the second ES. Similarly to the OS study design, women in the ES were administered annual mail-in questionnaires to obtain their follow-up data.

3.1.2 Additional Criteria for the Current Study

In addition to the enrollment criteria required by the WHI, for the current study, we applied the exclusion criteria of previous history of cancer (diagnosed prior to study enrollment) to minimize any potential bias, such as genetic predisposition or chemotherapy exposure. There were no other inclusion criteria applied to allow for a large sample size and to aid generalizability.

3.2 Research Design

Our study was a 1:4 nested case-control study within the CT and the OS of the WHI. For the analyses in Aims 1 and 2, the following participants were excluded from the original cohort of 161,808 for this analysis: 24,654 women who had a history of cancer at baseline, 7,148 women who had missing main exposure variables and 405
women with a new leukemia diagnosis (other than CLL/SLL) during the study. This resulted in a sample size of 129,601 subjects, of whom 328 were diagnosed with CLL/SLL during the WHI follow-up. Applying a random selection of four controls for each case, matched by age (5-year window) and race, provided 1312 controls (a total of 1640 subjects for statistical analyses).

To investigate pesticide exposures, the OS was used, as detailed pesticide exposures were only collected for women enrolled in this cohort. This study was also a nested case control study (1:4 case-control matching), done within the observational study arm of the WHI. The following participants were excluded from the original observational cohort of 93,676 for this analysis: 18,123 women who had a history of cancer at baseline, 12,506 women who had missing main exposure variables (either at baseline or at the year one follow-up form) and 359 women with a new leukemia diagnosis (other than CLL/SLL) during the study. This resulted in a sample size of 62,688 subjects, of whom 157 became confirmed CLL/SLL cases during the study. Applying a random selection of four controls for each case, matched by age and race, resulted in 628 controls and a total of 785 subjects for statistical analyses (after exclusions and matching).

3.3 Data Source

3.3.1 WHI Database

This study utilized data from the WHI. Three types of data were collected under
the WHI: self-reported (personal interview or questionnaire), clinical measurements, and outcome data. Self-reported data included demographic, medical history, diet, reproductive history, family history, and psychosocial and behavioral variables, and were obtained via standardized questionnaires adopted from similar studies. Clinical measurement data such as height, weight, waist/hip measures, blood pressure, functional status, and results from gynecologic exams were taken by certified WHI clinic staff using standard practice and recorded onto case report data collection forms specifically designed for the study. Blood specimens were banked and analyzed centrally, with some minor exceptions that were analyzed at local labs. Electrocardiogram and bone densitometry data were processed and read centrally. Clinical outcome data were initially self-reported. Outcomes important to the study (ex. cancer and cardiovascular disease) were then further confirmed via medical records to facilitate medical coding.156

As described above, the WHI is composed of two components: The Clinical Trials and the Observational Study. For a subset of the database, data collection forms and schedule of data collection were identical for the CT and OS, but for the majority of database each component has its own set of forms and schedule.

3.3.2 Clinical Trials Database

The Clinical Trials database of the WHI was compiled to allow for the evaluation of three types of interventions detailed above in Section 3.1. Subject level data availability for the CT database is highly heterogeneous as women could participate in any of seven combinations of trials.
Baseline data collected via the following forms are included in the CT database for all women participating in the CT: Eligibility Screen, Personal Information, Medical History, Reproductive History, Family History, Personal Habits, Thoughts & Feelings, Hormone Use Interview, Current Medications, Current Supplements, Food Questionnaire, Physical Measures, Waist/Hip Measures, and Blood Collection. In addition, subgroups of women had baseline data collected via the following forms in the CT database: Clinical Breast Exam, Mammogram, ECG, Bone Density, Functional Status, Pap Smear, Urine Collection, Cognitive Assessment, Pelvic Exam, Endometrial Aspiration, and HT washout.

The post baseline section of the CT database contains data extracted from the following forms for all women participating in the CT: Medical History Update (obtained semi-annually), Personal Habits Update (obtained every three years), Mammogram (obtained every two years), Daily Life (obtained at one year), Current Medications (obtained every three years), Current Supplements (obtained every three years), Physical Measures (obtained every two years), Waist/Hip Measures (obtained at one year), ECG (obtained at three, six, and nine years) and Blood Collection (obtained at one year). In addition, subgroups of women had post-baseline data collected via the following forms in the CT database: HT Manage/Safety Interview (obtained semi-annually), CaD Manage/Safety Interview (obtained semi-annually), Cognitive Assessment (obtained every three years), Food Questionnaire (obtained annually), Pelvic Exam (obtained annually), Endometrial Aspiration (obtained at three, six, and nine years), Clinical Breast Exam (obtained annually), Mammogram (obtained annually), Bone Density (obtained
every three years), Functional Status (obtained every three years), and Urine Collection (obtained at three, six, and nine years).

3.3.3 Observational Study Database

The Observational Study within the WHI was created in order to assess risk factors relative to diseases and mortality, as well as to serve as a control cohort for the CT component. The OS database therefore consists of longitudinal data on risk factor exposures, health status, disease incidence, and mortality.

Baseline data collected via the following forms are included in the OS database: Eligibility Screen, Personal Information, OS Questionnaire, Medical History, Reproductive History, Family History, Personal Habits, Thoughts & Feelings, Hormone Use Interview, Current Medications, Current Supplements, Food Questionnaire, Physical Measures, Waist/Hip Measures, and Blood Collection.

The post baseline section of the OS database contains data extracted from the following forms: Medical History Update (obtained annually), Daily Life (obtained at three years), Current Medications (obtained at three years), Current Supplements (obtained at three years), Food Questionnaire (obtained at three years), Physical Measures (obtained at three years), Waist/Hip Measures (obtained at three years), Blood Collection (obtained at three years), and Follow-up (obtained at three and six years).
3.4 Study Period and Population

The WHI study had an enrollment period spanning from October 1, 1993 to December 31, 1998 across 40 U.S. clinical centers. During this time, a total of 161,808 women were enrolled in the WHI study, with 93,646 taking part in the Observational Study and 61,132 in the Clinical Trials. The originally planned nine-year average follow-up period has been extended twice so far through two five-year long Extension Studies, the second of which is to end in 2015.

Participating women were between 50 and 79 years of age at the time of their enrollment, postmenopausal, and planned to reside in the area for three years or more. While the WHI study attempted to include 20% of the women from racial/ethnic minority groups, to achieve adequate sample sizes for statistical analyses within subsets according to minority groups, the study enrolled 17% of their participants from minority groups. With respect to the age distribution, there was an attempt to increase enrollment on the younger end of the menopausal age range (for biomarker studies) and on the older end also (for quality of life measured). Subsequently, the mean age of participants at baseline was 63 years (range: 49-81 years), 33% were from the <50 years-59 years age group, 45% from the 60 years-69 years age group, and 22% were >70 years. Regarding education level, nearly 40% of participants had a bachelor’s degree or higher and only about 5% did not obtain a high school diploma. At baseline, about half of the women were retired, the remaining still working, with fewer than 10% working as homemakers. The median yearly household income was around $40,000, with approximately 15% having a household income of less than $20,000 per year, and nearly 12% having a
household income of more than $100,000 per year. About 95% of the women had some
form of health insurance coverage, either through a private insurance company,
Medicare, or Medicaid. On average, the last time the enrolled women had visited a
doctor’s office was six months prior to enrollment to the WHI study and only 17.4% had
not been to a doctor during the one year period prior to their WHI study enrollment. 158

3.5 Data Elements

All data that were used in the current study were collected as part of the WHI
study –either via the CT or the OS- and were subject level. Selected baseline data
elements that were used from the WHI database are demographics, medical history,
reproductive history, and family history. In addition, a wide range of risk factor variables
collected longitudinally were used. Outcome measures extracted from the WHI database
were CLL/SLL incidence, including SLL cases. Finally, CT arm status was evaluated for its
association with CLL/SLL incidence [i.e., HT arm (intervention vs. control), DM arm
(intervention vs. control), and CaD arm (intervention vs. control)].

Furthermore, the OS questionnaire administered at screening ascertained some
additional risk factors, such as geographic residence history, secondhand smoking
exposure, early life exposures, details of physical activity, weight and weight cycling
history, and occupational exposures- these data were evaluated for the subset of women
participating in the OS.
3.5.1 Demographics

Participants in both the CT and the OS had a series of three baseline screening visits during which demographic data were obtained.

The following demographic variables collected at baseline were summarized and included in statistical analyses:

- Age [Continuous (years)]
- Age [50-54, 55-59, 60-69, 70-79 (years)]
- Self-reported race/ethnicity [American Indian or Alaskan Native, Asian or Pacific Islander, Black or African-American, Hispanic/Latino, White (not of Hispanic origin), Other];
- Region of U.S. (Northeast, South, Midwest, West);
- Education [Didn't go to school, Grade school (1-4 years), Grade school (5-8 years), Some high school (9-11 years), High school diploma or GED Vocational or training school, Some college or Associate Degree, College graduate or Baccalaureate Degree, Some post-graduate or professional, Master's Degree, Doctoral Degree (Ph.D, M.D., J.D., etc.)];
- Family income (Less than $10,000, $10,000 to $19,999, $20,000 to
$34,999, $35,000 to $49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 to $149,999, $150,000 or more);

• Current Occupation and Occupational History (Main Categories: Manager/Professional, Technical/Sales/Administrative, Service/Labor, Homemaker Only; Specific Sub-Categories: Farm Work, Worked with Hair Dyes)

3.5.2 Medical History

Medical history was obtained at screening for all women in the WHI as well as semi-annually for CT participants and annually for OS participants. Medical history variables were used to evaluate the history of cancer exclusion criteria for the primary analysis, as well as to evaluate them as potential risk factors for CLL. Selected medical history included history of any cancer [Yes/No].

3.5.3 Reproductive History

Reproductive history data were obtained at screening for all women in the WHI. Reproductive history variables were assessed as covariates in the analyses of HT use. The following reproductive history variables were analyzed:

• Age at menarche [continuous]
• History of menstrual irregularity and amenorrhea [periods: 
  Regular/Irregular; one year without period: Yes/No]

• History of pregnancy [number of term pregnancies, age at first term 
  pregnancy]

3.5.4 Personal Habits

Data on personal habits were collected at screening for all women enrolled in 
the WHI. In addition, for CT participants only, updated personal habits data on certain 
variables were obtained longitudinally, at every three years. For the primary analysis, 
the baseline personal habits data were evaluated. Additionally, if sample size allows, 
for CT participants longitudinal personal habits data available were used to evaluate 
changes from baseline personal habits relative to outcome.

The following personal habits variables were used:

• Smoking history [Never, Past, Current];

• Secondhand smoking exposure:
  
  o Years as a child lived with smoker [Never lived with a smoker, 
    \(<1, 1-4, 5-9, 10-18 \text{ (years)}\)\]
• Years as adult lived with smoker [Never lived with a smoker, <1, 1-4, 5-9, 10-19, 20-29, 30-39, 40+ (years)]

• Years worked with smoker [Never worked with a smoker, <1, 1-4, 5-9, 10-19, 20-29, 30-39, 40+ (years)]

• Alcohol intake [Never drinker, Past drinker, Current drinker (<1 drink per month, <1 drink per week, 1-<7 drink per week, 7+ drinks per week)]

• Coffee consumption:
  
  o Coffee or tea consumption [Yes/No] (Note: this variable was collected in the Food Questionnaire at screening for all women and also annually for a subset of CT participants- only the screening data were used in the current study);

  o Coffee consumption [Cups of regular coffee per day: None, 1 cup, 2-3 cups, 4-5 cups, 6 or more cups] (Note: this variable was collected in the OS Questionnaire at screening – only the subset of OS participants had this factor evaluated for outcome)

portion]/glycemic load based on total carbohydrates [no unit] (Note: continuous computed variables from the Food Questionnaire, only screening data were used);

• Physical activity:

  o Total episodes per week of recreational physical activity (includes walking, mild, moderate and strenuous physical activity)[continuous];

  o Categorical variable of episodes per week of moderate and strenuous recreational physical activity of ≥ 20 minutes duration (includes walking fairly fast or very fast, moderate physical activity and strenuous physical activity, MET ≥ 4.0) [1= no activity; 2=some activity of limited duration, frequency or intensity; 3=moderate or strenuous activity of at least 20 minutes duration and 2 to less than 4 times per week; 4=moderate or strenuous activity of at least 20 minutes duration and 4 or more times per week];

  o Total minutes of recreational physical activity per week (includes walking, mild, moderate and strenuous physical activity) [continuous];
- Hours per week (kcal/week per kg) energy expenditure from walking [continuous];

- Minutes per week of physical activity from yardwork [continuous].

- Weight change history (note: these variables are collected on the OS Questionnaire at screening – only the subset of OS participants had this factor evaluated for outcome):
  - Maximum adult weight (lbs) [continuous];
  - Minimum adult weight (lbs) [continuous];
  - Times lost ≥ 50 lbs [None, 1-2, 3-4, 5-6, 7 or more (times)];
  - Years at being within 10 lbs of current weight (years) [continuous].

### 3.5.5 Family History

Based on previous findings, the following family history variables collected at screening were used to evaluate their association with leukemia incidence:

- Relatives' history of cancers
Female relatives (mother, full-blooded sisters, daughters, grandmothers) ever have cancer [Yes, No];

Female relatives (mother, full-blooded sisters, daughters, grandmothers) ever have breast cancer [Yes (<45 years old/45 years or older), No];

Male relatives (father, full-blooded brothers, sons) ever have cancer [Yes, No];

Any close relative ever have cancer [Yes, No].

3.5.6 Miscellaneous Repeated Measures Variables

The following additional variables measured at baseline and longitudinally (on an annual basis) were evaluated for their relationship with CLL, based on their established or suspected association; only exposure data collected prior to leukemia diagnosis date were used for cases and the exposure data were collected according to the same time frame from baseline for the corresponding matched controls:

- Weight, height

  - Body Mass Index (BMI) were a derived variable, according to the following formula: \( \text{BMI} = \frac{\text{weight (kg)}}{[\text{height (m)}]^2} \)
BMI [Underweight (<18.5), Normal (18.5–24.9), Overweight (25.0–29.9), Obesity I (30.0–34.9), Obesity II (35.0–39.9), Obesity III (≥40)]

- Insecticide/Pesticide use (Note: this variable was not collected at screening but was collected annually in the OS Follow-up Questionnaire, therefore only OS women were evaluated for this risk factor):
  - Location of exposure to insecticides [No, Yes, at work only, Yes, at home or leisure only, Yes, both at work and at home or leisure];
  - Mixed insecticides [Yes, No];
  - Sprayed or applied insecticides [Yes, No];
  - Lawn service applied insecticides [Yes, No];
  - Commercial service applied insecticides [Yes, No];
  - Other exposure to insecticides [Yes, No].

- Medications (collected at screening, every three years for CT participants, and at three years for OS participants)
  - Medication National Drug Code (NDC) (note: only codes for
NSAIDs were evaluated:\(^{159}\);

- Duration (in years) of medication use (since the age of 20 years).

- Complete blood count
  - Hemoglobin (g/dL) [continuous];
  - Platelets (10\(^3\)/µL) [continuous];
  - White Blood Cells (10\(^3\)/µL) [continuous].

### 3.5.7 Outcome

Self-reported health outcomes including cancer were collected on all women in a longitudinal manner. OS participants completed these health assessment questionnaires annually and CT participants semi-annually. Cancer cases were pathologically confirmed and coded according to the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) guidelines. This coding allows for the distinction of specific leukemia subtypes, such as Chronic Lymphocytic Leukemia (CLL). In addition, incident Small Lymphocytic Lymphoma (SLL) cases were coded as CLL, according the most recent WHO guidelines. Furthermore, survival status was collected periodically through follow-up and via regular searches in the National Death Index. Clinical adjudicators determined causes of death by medical record and death certificate review.
Clinical outcome measures extracted will include the following variables:

- Diagnosis of CLL (including SLL) [Self-reported by any of the following: pathology/cytology report, hospital Face Sheet with ICD-9-CM codes, operative report, hospital discharge summary, outpatient, day surgery, or short stay recode]

- Confirmation of CLL/SLL status
  - microscopically confirmed [Yes, No]

Access to the WHI database on the WHI Study Operations website has been granted via the WHI Data Distribution Agreement. A copy of this document is enclosed in the Appendix. In addition, since this is was a secondary data analysis and was not research involving human subjects, the Ohio State University IRB office waived the need for review.

3.6 Statistical Analyses

3.6.1 Sample Size Calculation

There were 450 confirmed CLL cases (n= 330 CLL, n= 120 SLL) diagnosed in the WHI study (CT+OS) during the period of October 1, 1993- January 31, 2014. Using a
matched, case-control study design with four controls per case (n = 1800 controls) allowed us to conduct a study where normally (with 1:1 matching) n = 720 cases would be required. Cases were matched to controls by age and race. Controls were selected from the WHI study and did not have had any cancer diagnoses at any point in their lives. The total sample size resulting from our 1:4 case-control design is n = 2250 (n = 450 cases, n = 1800 controls), which was adequate to evaluate already established and potential novel risk factors for their association with CLL/SLL incidence. Assuming a conservative estimate of correlation for exposure of 0.15 between matched cases and controls, with a probability of exposure among controls of 0.1 and an odds ratio (OR) of 1.6, we had approximately 80% power to detect a difference between cases and controls, at α = 0.05. The total reduction in sample size is nearly 40% by using a 1:4 rather than a 1:1 matching design. As found in previous epidemiologic studies, due to the weaker effect of some of the exposure variables evaluated, the ORs may be lower than 1.6 and closer to the 1.2-1.3 range. Although the p-values associated with lower ORs were larger they could nevertheless suggest trends of statistical association.

Depending on the number of women with available main exposure data for each aim, the sample size calculations were adjusted accordingly as detailed under each aim below in chapters 4, 5, and 6.

3.6.2 Descriptive Analyses

Descriptive analyses were reported by specific aim (see Chapter 1). In general, continuous baseline measurements were summarized by sample size, mean, and standard
error values. Categorical baseline variables were summarized by number and percentage of subjects in each category. Summary tables were provided for the overall study population as well as by outcome groups [i.e., CLL/SLL cases vs. controls (leukemia-free cases)].

In addition, descriptive analyses on outcome data were provided as follows: the number of CLL/SLL cases by age group, race, smoking status, and region.

3.6.3 Analysis of Personal Habits as Risk Factors for CLL/SLL (Specific Aim 1)

Baseline characteristics were summarized for cases and controls displaying mean and standard error for continuous variables and frequencies with percentages for categorical variables. To evaluate if a risk factor was associated with case-control status (for variables not used in the matching), we used univariable conditional logistic regression models with CLL/SLL case-control status as the dependent variable and the risk factor as the independent variable (either categorical or continuous). P-values from the likelihood ratio test are reported.161

To evaluate important personal habits risk factors for CLL/SLL incidence in the presence of other important risk factors, multivariable conditional logistic regression analysis was conducted. Odds ratios (ORs) and corresponding 95% CIs and p-values were calculated for the matched-pair cohort data and ORs were used to estimate risk ratios (due to CLL/SLL being a rare disease).162 Interactions between personal habits factors and other risk factors (i.e., study part (CT, OS), obesity, smoking status) were
evaluated by including a multiplicative interaction term in the regression model to evaluate any effect modification.

Analyses were performed using SAS 9.3 (SAS Institute, Cary, NC).

3.6.4 Analysis of Hormone Exposures as Risk Factor for CLL/SLL (Specific Aim 2)

Baseline characteristics were summarized for cases and controls displaying mean and standard error for continuous variables and frequencies with percentages for categorical variables. To compare the groups for variables not used in the matching, univariable conditional logistic regression models were generated and the likelihood-ratio p-values were reported. Current HT use, obtained at the time of randomization (for HT trial participants) and at the three-year follow-up (for all other women not part of the HT trial), were summarized.

Univariable conditional logistic regression models estimated CLL/SLL risk and past OC use, past HT use, current HT use and other important risk factors. To evaluate important hormone use related risk factors for CLL/SLL incidence in the presence of other important risk factors, multivariable conditional logistic regression analysis was conducted. Odds ratios (ORs) and corresponding 95% confidence intervals and p-values were calculated for the matched-pair cohort data and ORs were used to estimate risk ratios. Interactions between EH use and other risk factors (i.e., obesity, smoking status) were evaluated by including a multiplicative interaction term in the regression
analyses were performed using SAS 9.3 (SAS Institute, Cary NC).

3.6.5 Analysis of Pesticide Exposures as Risk Factors for CLL/SLL (Specific Aim 3)

Baseline characteristics were summarized for cases and controls displaying mean and standard error for continuous variables and frequencies and percentages for categorical variables. To compare the groups for variables not used in the matching, we used univariable conditional logistic regression models. Past pesticide exposure variables obtained at the one year follow-up were summarized similarly, as well as the surrogate pesticide exposure variable “Ever lived or worked on a farm?”. Cumulative pesticide exposure was divided into quartiles and summarized accordingly. Univariable conditional logistic regression models were generated for CLL/SLL incidence for all collected pesticide use variables and other important risk factors. To evaluate important pesticide use related risk factors for CLL/SLL incidence in the presence of other important risk factors, multivariable analysis was conducted. Interactions between pesticide use and other risk factors (i.e., obesity, smoking status) were evaluated by including a multiplicative interaction term in the regression model.

Conditional logistic regression models adjusted for important risk factors for the matched-pair cohorts, were used to estimate risk ratios of pesticide exposure for CLL/SLL incidence. In addition, all obtained risk ratio estimates, with their corresponding 95% confidence intervals and p-values, were calculated and reported. There were no adjustments made for multiple testing since only a few planned
comparisons were made in this analysis. Analyses were performed using SAS 9.3 (SAS Institute, Cary NC).

### 3.6.6 Summary of Statistical Analyses Methods

In general, similar statistical considerations and methods were used throughout the three aims of our study. The design was a nested case-control study in all three aims, and controls were randomly matched by age and race to cases. In aims 1 and 2, both the CT and OS parts of the WHI were utilized for the studies. In aim 3, as detailed pesticide data was only captured in the OS part, the study was based on the OS cohort only. All descriptive analyses used the same reporting format of presenting the mean and standard error for continuous variables and the frequency and percentage for categorical variables. All baseline variable comparisons were made by generating univariable conditional logistic regression models to evaluate associations with the dichotomous outcome of case/control status (i.e., CLL/SLL risk), and the p-values from the likelihood ratio tests were reported.

Multivariable modeling was conducted to fully investigate all three aims. In each setting the primary exposure variables of interest were identified based on their univariable significance level, which also included evaluation of potential interactions. Variables with a p-value of .10 or less were considered for inclusion in the full multivariable model and a backward elimination procedure was used to arrive at the final models. Smoking status (ever vs. never), the region of the U.S., and study part (CT, OS) were retained in all final models, irrespective of their statistical significance.
Chapter 4: Personal Habits and their Association with Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma Incidence

4.1 Background

Although leukemia is a relatively rare form of cancer in adults (approximately 3% of all cancers), there are about 240,000 new cases diagnosed annually worldwide (43,800 in the U.S.) and approximately 200,000 deaths associated with leukemia (23,300 in the U.S.) each year.\textsuperscript{7,8} Leukemia is ranked fifth in person-years of life lost due to cancer, directly behind breast and pancreatic cancer.\textsuperscript{9} According to the National Cancer Institute’s Surveillance, Epidemiology and End Results (SEER) Program, during recent decades (between 1975 and 2010), there has been a trend of increasing incidence of leukemia among women— with no clear explanation to support this trend.

In industrialized countries, Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma (CLL/SLL) is the most common type of adult leukemia, and is associated with aging, with causes largely unknown. Despite being a prevalent type of leukemia, CLL/SLL is a relatively rare form of cancer (approximately 1% of all cancers). There are still about 15,720 new CLL/SLL cases diagnosed each year in the U.S – mainly in older adults (the lifetime risk of CLL/SLL is 0.52%, median age at diagnosis: 70 years).\textsuperscript{7,10} Since CLL/SLL is twice as common among males compared to females,
studying women may reveal factors, unique to women, that are potentially protective against CLL/SLL.

As CLL/SLL typically presents later in life, following several decades of DNA damage from endogenous and exogenous sources, there is a need to recognize the multiple types of exposures that may contribute to CLL/SLL risk, including personal habits. Currently, very little is known about how personal habit exposures, such as diet and exercise, potentially relate to CLL/SLL risk. This is likely to be, in part, attributable to the difficulty in reliably assessing these exposures, as reflected by the few studies that have assessed these exposures for CLL/SLL risk without finding consistent results. In addition, most studies tend to be underpowered due to the relative rarity of CLL/SLL, which contributes to the difficulty in getting consistency across studies, especially when evaluating exposures with weaker associations.

The largest study to date evaluating diet and CLL/SLL incidence has been conducted by Tsai et al.\textsuperscript{63} This was a meta-analysis of two US based national cohort studies and included 338 women with incident CLL/SLL. The authors did not find any association between dietary factors and CLL/SLL incidence when both sexes were considered and since men comprised most of the CLL/SLL cases in this analysis, it is not clear what the results pertaining to women only would be. A recently published case-control study from Spain, also conducted on men and women combined, found higher fruit intake to be related to increased CLL/SLL incidence.\textsuperscript{163} However, this finding was not supported by any biologic rationale and the results may be attributable to higher
pesticide exposures from increased fruit intake.

With respect to drinking habits, alcohol, tea, and coffee are the key areas of interest due to their use by a substantial proportion of women and potential association with cancer risk. Morton et al. conducted a large meta-analysis of alcohol use and NHL (including 991 CLL/SLL cases) risk and found alcohol use to be protective against NHL overall, but when considering CLL/SLL as a subset, these results did not persist.\(^{164}\) Ji et al. studied subjects with alcohol use disorder in Sweden and found alcohol to be protective against hematological malignancies, but they did not provide results for CLL/SLL separately.\(^ {165}\) On the other hand, Parodi et al. in an Italian case-control study, did not identify any associations between alcohol consumption and the risk of CLL/SLL, perhaps because they only studied 71 CLL/SLL cases.\(^ {166}\)

Relative to coffee and tea consumption, the study by Parodi et al. found black tea to be protective, but did not find an association with coffee drinking. An earlier Italian study similarly did not identify any associations with coffee drinking and NHL risk.\(^ {167}\) In a hospital based case-control study conducted in China by Zhang et al., green tea was identified as a protective risk factor for all adult leukemia, however this study only included 3 CLL/SLL cases.\(^ {64}\) Another Asian study conducted by Balasubramaniam et al. in India did not find any associations between tea drinking, which is very common in India, and leukemia risk but did for coffee drinking (less common than tea drinking in India), showing a 50% reduction in NHL risk among coffee drinkers.\(^ {168}\) There are certain biological mechanisms that support the protective role of coffee and tea, as both contain
polyphenols that have the potential to inhibit DNA methylation. In addition, coffee has been found to be weakly estrogenic. Coffee contains phytoestrogens some of which, such as trigonelline, can activate estrogen receptors and others, such as enterolactone which has been shown to lower the risk of estrogen dependent cancers.

Exercise, a known immune system modifier, is also an important candidate risk factor for CLL/SLL development. Nevertheless, there is no consensus relative to its impact based on the epidemiologic investigations to date: some have found associations and some did not. In the first study to investigate this risk factor, Cerhan et al. studied a small set (n=30) of CLL/SLL cases as part of the Iowa Women’s Health Study and did not find any association between physical activity levels and CLL/SLL risk. In a later study, Cerhan et al. studying a different cohort of NHL cases, found physical activity to be protective against NHL risk but this effect did not remain after adjusting for body mass index (BMI) and alcohol consumption and they did not present their data relative to CLL/SLL. Another study of NHL, which included CLL/SLL cases (n=124) from the California Teachers Study, Lu et al. could not find any effect of physical activity on CLL/SLL risk. Pan et al. conducted a population based case-control study of NHL in Canada and found physical activity to be protective against NHL risk in both men and women, but they did not conduct separate analyses for CLL/SLL.

In order to recommend guidance to the public on healthy choices regarding body weight, exercise, nutrition, and alcohol use, the American Cancer Society (ACS) periodically publishes Nutrition and Physical Activity Guidelines, which was last updated
in 2012. Although these guidelines are targeted towards reducing cancer incidence and mortality due to cancer, they largely overlap with other healthy living recommendations, such as ones by the American Heart Association and, therefore, are also intended for general disease and mortality reduction. The ACS Nutrition and Physical Activity Guidelines incorporate four personal habit associated components (body weight, physical activity, diet, and alcohol consumption) into one score on a nine-point scale. The higher the ACS score, the healthier the behavior, whereas lower scores suggest less healthy personal habits. A recently published study on the Women’s Health Initiative (WHI) observational study (OS) cohort evaluated the ACS score on cancer incidence (and cancer specific mortality) for the most frequent cancers in women (i.e., breast, colorectal, endometrial, ovarian, and lung), however, it did not evaluate rarer cancers, such as CLL/SLL separately, only as part of a miscellaneous group of less common cancers. The highest ACS scores compared with the lowest associated with lower breast, colorectal, and endometrial cancer incidence, but did not associate with ovarian, lung, or the miscellaneous group of cancers. Therefore, it is unclear if there are any associations of the ACS score with CLL/SLL risk.

We investigated the association between personal habit exposures and CLL/SLL using data from the WHI, a large prospective study of post-menopausal women. We evaluated the association between potentially important nutritional and drinking exposures, body weight, and physical exercise and the risk of CLL/SLL. We also investigated whether there was any relationship between a combined score evaluating healthy lifestyle (i.e., ACS Nutrition and Physical Activity Guidelines score) and
4.2 Methods

Study Design

The WHI was designed to address the major causes of morbidity and mortality in postmenopausal women and includes four clinical trials (CTs) and an observational study (OS). Details of the scientific rationale, eligibility requirements, and baseline participant characteristics of the WHI have been published elsewhere. Briefly, a total of 161,808 women, 50–79 years of age, were recruited at 40 clinical centers throughout the United States between September 1, 1993 and December 31, 1998. The four CTs were: two hormone therapy (HT) trials (27,347 women), a dietary modification trial (DM; 48,835 women), and a calcium/vitamin D supplementation trial (CaD; 36,282 women). With respect to the two HT trials, one study investigated estrogen-alone (E-alone) in post-menopausal women without a uterus with the experimental arm taking a daily dose of estrogen in the form of conjugated equine estrogen (CEE) for an average of about 6 years. The other study looked at estrogen plus progestin (E+P) in post-menopausal women who still had their uterus; women on the experimental arm of this study took a daily dose of CEE plus a progestin (medroxyprogesterone acetate) for an average of about 5 years. Both HT trials had placebo control arms. Participants in the OS included 93,676 women who were screened for the clinical trials but proved to be ineligible or unwilling to participate, or who were recruited through a direct invitation for the OS. A large percentage (72.8%) of women in
the OS took HT (E-alone or E+P) while enrolled in the study. All WHI participants signed informed consent and were followed prospectively. The WHI study was overseen by institutional review boards at all 40 clinical centers and at the coordinating center, and by a study-wide data-safety monitoring board.

This study was a 1:4 nested case-control study within the CT and the OS of the WHI. The following participants were excluded from the original cohort of 161,808 for this analysis: 24,654 women who had a history of cancer at baseline, 7,148 women who had missing main exposure variables and 405 women with a new leukemia diagnosis (other than CLL/SLL) during the study. This resulted in a sample size of 129,601 subjects, of whom 328 were diagnosed with CLL/SLL during the WHI follow-up. Applying a random selection of four controls for each case, matched by age (5-year window) and race, provided 1312 controls (a total of 1640 subjects for statistical analyses).

**Measures**

*Dietary habits*

Diet was measured at baseline using a validated, self-administered food frequency questionnaire that was developed by the WHI. This questionnaire was adapted from the Health Habits and Lifestyle Questionnaire, which was based on the Second National Health and Nutrition Examination Survey (NHANES II) study. Both single item (e.g., red meat intake) and composite (e.g., saturated fat) questions were included on the questionnaire.
**Drinking habits**

Data on drinking habits, including alcohol and coffee consumption were collected at baseline by self-administered questionnaire. Relative to alcohol drinking, multiple questions were asked, including ones about daily amount of beer, wine, and liquor consumption. For coffee drinking, women were asked if they drank coffee usually each day, not including decaffeinated coffee.

**Physical activity**

Data on exercise habits was evaluated for all women in the study at baseline. Several measures were assessed, including times and duration for moderate and hard exercise by week. In addition, metabolic equivalent (MET) hours per week were computed to obtain intensity along with energy expenditure.

**ACS Nutrition and Physical Activity Cancer Prevention Guidelines score**

To evaluate an overall personal habit summary measure, in addition to the single component measures detailed above, we calculated the nine-level (0-8) ACS Nutrition and Physical Activity Cancer Prevention Guidelines score for the women in our study. This score is computed using diet, BMI, physical activity, and alcohol use data, as outlined in the published guidelines. The diet, physical activity, and alcohol habits were assessed as described above. BMI was calculated using body weight and height measured by trained staff at the baseline clinic visit.

**Follow-up and ascertainment of cases**
According to the World Health Organization (WHO), since 2008, SLL and CLL have been considered one disease and one entity for disease classification, as malignant cells in both diagnoses exhibit the same immunophenotype. In our study, we followed the most current guidelines (WHO 2008) that include SLL cases along with CLL.

Incident CLL/SLL cases were identified by self-administered questionnaires (administered annually in the WHI CT after 2005, and annually in the WHI OS throughout the study), with all cases confirmed by medical record review. All CLL/SLL cases then were coded centrally in accordance with the Surveillance Epidemiology and End Results (SEER) coding guidelines. For identification of cases, participants were followed up to CLL/SLL diagnosis, date of death, loss to follow-up, or end of WHI CT or OS follow-up, whichever occurred first.

**Statistical analysis**

Baseline characteristics were summarized for cases and controls displaying mean and standard error for continuous variables and frequencies with percentages for categorical variables. To determine whether a risk factor was associated with case-control status (for variables not used in the matching), we used univariable conditional logistic regression models with CLL/SLL case-control status as the dependent variable and the risk factor as the independent variable (either categorical or continuous). P-values from the likelihood ratio test are reported.161

To evaluate important personal habits risk factors for CLL/SLL incidence in the presence of other important risk factors, multivariable conditional logistic regression
analysis was conducted. Odds ratios (ORs) and corresponding 95% CIs and p-values were calculated for the matched-pair cohort data and ORs were used as a surrogate for relative risks (due to CLL/SLL being a rare disease). Interactions between personal habits factors and other risk factors (i.e., study part (CT, OS), obesity, smoking status) were included as multiplicative interaction terms in the regression model to assess the role variables may play as effect modifications.

4.3 Results

Descriptive Data Analysis

A total of 328 CLL/SLL cases meeting study inclusion criteria (i.e., no previous history of cancer, non-missing main exposure variables) were identified, during a mean follow-up period of 13.8 years. Baseline characteristics according to case - control status are shown in Table 4.1. Since age and race were both used in matching controls to cases, these variables were evenly distributed between the groups by definition. With respect to other important baseline characteristics, body mass index (BMI) measured at the baseline was lower for cases (P=.02). In addition, cases were more often college graduates compared to controls (P=.05). Smoking status (i.e., “Smoked at least 100 cigarettes ever”), region of the U.S. they resided in, and all other demographic factors were similar between cases and controls. With respect to physical activity, cases and controls exercised similar amounts (P=.24).

Relative to dietary habits, the majority of items evaluated were consumed in
similar amounts among cases and controls, with dark fish consumption being the only statistically different factor, suggesting that cases consumed it in larger amounts ($P=.02$), however, due to the low frequency of consumption of this food overall in our cohort, the effect size for this comparison is small to draw meaningful conclusions.

Overall alcohol consumption was similarly low between cases and controls with around a third of a serving/day ($P=.66$) and when considering the type of alcohol, there appeared to be a slightly higher consumption of beer among controls, compared to cases (0.04 vs. 0.03 beer servings/day, $P=.10$), suggesting a possible protective effect of beer against CLL/SLL bearing in mind the small effect size. There were no differences for wine or liquor consumption between the groups.

Coffee consumption and the amount of caffeine (from all beverage sources) did not differ between cases and controls ($P=.33$ and $P=.42$, respectively). The 9-level ASC score was similar among cases and controls ($P=.15$; 0-2 v 3 v 4 v 5 v 6-8), and this analysis did not suggest that higher, more desirable scores offered protection against CLL/SLL among women.

**Multivariable Model Evaluating CLL/SLL Risk**

To further explore any relationship of the above specified personal habit risk factors, we evaluated effect modification due to study arm for each factor, as women tended to be different at baseline between the CT and the OS (Table 4.2) for several characteristics. Specifically, women on the OS had more frequently used
estrogen+progestin therapy in the past (P<.001). Also women enrolled on the OS tended to be generally healthier and, therefore, had higher ASC scores (P<.001), compared to CT women. In addition, to the components that go into the ASC score, the only other personal habit parameter that seemed different at baseline between the CT and OS women was that OS women consumed less coffee (P=.11) and caffeine (P=.01). Upon multivariable conditional logistic regression modeling (Table 4.3), although coffee consumption did not show a significant interaction with study part (P=.22), coffee drinkers in the CT arm had their risk of CLL reduced by 27% (P=.09), whereas no effect of coffee was observed for OS women (P=.99). Furthermore, our multivariable model showed that past oral contraceptive use (OR=0.74, 95% CI: 0.56, 0.96; P=.03; CT+OS) and obesity (OR=0.71, 95% CI: 0.53, 0.94; P=.20; CT+OS) reduced CLL/SLL risk, whereas past estrogen use (OR=1.32, 95% CI:1.02, 1.71; P=.04; CT+OS) was an adverse risk factor. Neither region of U.S. or smoking were significantly related to CLL/SLL risk.

### 4.4 Discussion

In this nested case-control study, designed using a large prospective cohort of postmenopausal women, we found that a subset of women we studied (i.e., those on the CT of WHI) who were regular coffee drinkers had a reduced risk of developing CLL/SLL. This finding remained after controlling for a set of significant CLL/SLL risk factors, past OC use, past estrogen-alone therapy, and obesity. In addition to identifying coffee as a protective risk factor, past OC use and being obese were also protective, suggesting that increased estrogen levels (both from exogenous and endogenous sources)
are beneficial in preventing CLL/SLL. On the contrary, women who took estrogen alone therapy in the past were at an increased risk for CLL/SLL, likely due to being deficient in estrogen after bilateral oophorectomies.

Coffee has several bioactive components of interest for various health indications, such as cancer and type 2 diabetes. Some of these compounds have been identified to have specific mechanisms by which they potentially influence leukemogenesis. Caffeine has been found to suppress B-cell CLL/lymphoma (BCL2) gene expression in laboratory studies. Also, kahweol, a diterpene found in coffee only, has been found to inhibit tumor cell growth by modulation of several parts of leukemia cell apoptosis. Coffee and caffeine consumption also appears to modulate estrogen receptor activity. In a study of women who were given various amounts of coffee and caffeine, those with the highest levels consumed, had the highest estrogen levels in their urine. Considering that there is a connection between higher estrogen levels offering protection against CLL/SLL, this mechanism can explain our finding.

As all of the other factors that remained significant in our final multivariable model have estrogenic activity (i.e., hormone therapy and obesity) there is a strong possibility that the underlying mechanism for the impact of coffee in our study is also due to its estrogenic potential. As we show in Table 4.3, there were multiple important baseline differences between women on the CT compared to those on the OS. Among these, the most relevant to our study was that women on the OS had a history of more frequent estrogen+progestin therapy use compared to CT participants. We hypothesize that this hormone therapy use led to increased estrogen in the OS women while CT
women still had some unmet estrogen needs which then could be compensated for by consuming phytoestrogens from diet, such as by coffee drinking. In addition, since women on the CT tended to be more often regular coffee drinkers compared to women on the OS, there is possibly more potential to pick up an effect in this group.

Our study did not identify any of the other drinking, dietary, or exercise measures evaluated to associate with CLL/SLL risk. The ACS score, which is a composite measure of several important behavioral exposures, also did not associate with the risk of CLL/SLL. One explanation for the lack of impact of the ACS score is that one of its four main components is BMI, and obesity is, in fact, protective against CLL/SLL. Moreover, exercise, another one of the remaining three ACS score components, is inversely related to BMI, therefore it has an opposite association.

This study has multiple strengths including 1) the nested case-control design that enabled us to efficiently assess the risk of CLL/SLL which is a rare cancer and is therefore difficult to study in a cohort setting; 2) its focus on women relative to a wide array of personal habit exposures and their association with CLL/SLL in an adequately powered study; 3) confirmed CLL/SLL diagnosis; and 4) use of a well-designed and validated food questionnaire. However, there are some possible limitations to our study, such as our main exposure variables were self-reported and some level of recall bias and misclassification was possible. In addition, we were only able to evaluate exposures at baseline, with the assumptions that most personal habits are constant over time. Also, as CLL/SLL take decades to develop, from our study it is not clear what the timing and dose of an exposure needs to be in order to have an association with CLL/SLL risk. Finally,
pesticide exposure, a known adverse risk factor for CLL/SLL was not assessed in this current study, due to not being assessed for women on the CT. \textsuperscript{188}

In conclusion, postmenopausal women on the CT of WHI who had a history of habitual daily coffee drinking had a reduced risk for CLL/SLL, by 27%. Our study did not identify any other drinking, dietary, or exercise habits to be associated with CLL/SLL incidence, neither by single components, nor by the ACS score. Our multivariable model may explain the lower incidence of CLL/SLL among women compared to men as all significant findings are estrogen related. Biological studies are needed to support our findings, in particular relative to the mechanisms of coffee on CLL/SLL development.
### Table 4.1. Baseline Characteristics and WHI Study Participation by Control and Case Status

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=1312)</th>
<th>Cases (n=328)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>62.9±0.2</td>
<td>63.2±0.4</td>
<td>Matched</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American (%)</td>
<td>80 (6.1)</td>
<td>20 (6.1)</td>
<td></td>
</tr>
<tr>
<td>White (%)</td>
<td>1188 (90.6)</td>
<td>297 (90.6)</td>
<td></td>
</tr>
<tr>
<td>Other (%)</td>
<td>44 (3.3)</td>
<td>11 (3.3)</td>
<td></td>
</tr>
<tr>
<td>U.S. Region</td>
<td></td>
<td></td>
<td>.54</td>
</tr>
<tr>
<td>Northeast (%)</td>
<td>310 (23.6)</td>
<td>86 (26.2)</td>
<td></td>
</tr>
<tr>
<td>South (%)</td>
<td>320 (24.4)</td>
<td>86 (26.2)</td>
<td></td>
</tr>
<tr>
<td>Midwest (%)</td>
<td>291 (22.2)</td>
<td>67 (20.4)</td>
<td></td>
</tr>
<tr>
<td>West (%)</td>
<td>391 (29.8)</td>
<td>89 (27.1)</td>
<td></td>
</tr>
<tr>
<td>Body mass index (BMI), kg/m²</td>
<td>28.1±0.2</td>
<td>27.3±0.4</td>
<td>.02</td>
</tr>
<tr>
<td>Body mass index category*</td>
<td></td>
<td></td>
<td>.03*</td>
</tr>
<tr>
<td>Underweight (&lt; 18.5) (%)</td>
<td>12 (0.9)</td>
<td>2 (0.6)</td>
<td></td>
</tr>
<tr>
<td>Normal (18.5 - 24.9) (%)</td>
<td>426 (32.5)</td>
<td>116 (35.4)</td>
<td></td>
</tr>
<tr>
<td>Overweight (25.0 - 29.9) (%)</td>
<td>472 (36.0)</td>
<td>129 (39.3)</td>
<td></td>
</tr>
<tr>
<td>Obesity I (30.0 - 34.9) (%)</td>
<td>251 (19.1)</td>
<td>56 (17.1)</td>
<td></td>
</tr>
<tr>
<td>Obesity II (35.0 - 39.9) (%)</td>
<td>98 (7.5)</td>
<td>18 (5.5)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 4.1. Baseline Characteristics and WHI Study Participation by Control and Case Status (continued)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Case</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extreme Obesity III (≥40) (%)</td>
<td>53 (4.0)</td>
<td>7 (2.1)</td>
<td></td>
</tr>
<tr>
<td>Physical activity* (MET[^b] hours/week)</td>
<td>6.29±0.2</td>
<td>7.01±0.56</td>
<td>.24</td>
</tr>
</tbody>
</table>

Diet

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Case</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fruit and vegetables*, servings/day</td>
<td>4.09±0.07</td>
<td>4.06±0.11</td>
<td>.82</td>
</tr>
<tr>
<td>Vegetables, servings/day</td>
<td>2.20±0.04</td>
<td>2.21±0.07</td>
<td>.92</td>
</tr>
<tr>
<td>Fruits, servings/day</td>
<td>1.89±0.03</td>
<td>1.85±0.07</td>
<td>.61</td>
</tr>
<tr>
<td>Total carotenoids*, mcg/day</td>
<td>10,905±168</td>
<td>10,606±307</td>
<td>.42</td>
</tr>
<tr>
<td>Red and processed meat*, servings/day</td>
<td>0.94±0.02</td>
<td>0.91±0.04</td>
<td>.55</td>
</tr>
<tr>
<td>Fish, servings/day</td>
<td>0.26±0.01</td>
<td>0.28±0.01</td>
<td>.14</td>
</tr>
<tr>
<td>Dark fish, servings/day</td>
<td>0.03±0.00</td>
<td>0.05±0.00</td>
<td>.02</td>
</tr>
<tr>
<td>White fish, servings/day</td>
<td>0.06±0.00</td>
<td>0.07±0.00</td>
<td>.56</td>
</tr>
<tr>
<td>Soy, servings/day</td>
<td>0.02±0.00</td>
<td>0.02±0.01</td>
<td>.90</td>
</tr>
<tr>
<td>Dairy, servings/day</td>
<td>1.86±0.04</td>
<td>1.89±0.07</td>
<td>.71</td>
</tr>
<tr>
<td>Whole grains, servings/day</td>
<td>1.20±0.02</td>
<td>1.20±0.05</td>
<td>.98</td>
</tr>
<tr>
<td>Proportion of grains as whole grains*</td>
<td>0.27±0.00</td>
<td>0.26±0.01</td>
<td>.10</td>
</tr>
<tr>
<td>Fats/oils, total g/day</td>
<td>62.17±1.04</td>
<td>61.70±1.64</td>
<td>.83</td>
</tr>
<tr>
<td>Saturated fats, g/day</td>
<td>21.05±0.38</td>
<td>20.48±0.59</td>
<td>.47</td>
</tr>
<tr>
<td>Monounsaturated fats, g/day</td>
<td>23.57±0.40</td>
<td>23.43±0.64</td>
<td>.87</td>
</tr>
<tr>
<td>Polyunsaturated fats, g/day</td>
<td>12.64±0.21</td>
<td>12.95±0.35</td>
<td>.50</td>
</tr>
<tr>
<td>Trans fats, g/day</td>
<td>4.42±0.09</td>
<td>4.51±0.16</td>
<td>.62</td>
</tr>
<tr>
<td>Alcohol*, drinks/day</td>
<td>0.34±0.02</td>
<td>0.36±0.04</td>
<td>.66</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th></th>
<th>Control (%)</th>
<th>Case (%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beer, drinks/day</td>
<td>0.04±0.01</td>
<td>0.03±0.01</td>
<td>.10</td>
</tr>
<tr>
<td>Wine, drinks/day</td>
<td>0.20±0.01</td>
<td>0.21±0.03</td>
<td>.71</td>
</tr>
<tr>
<td>Liquor, drinks/day</td>
<td>0.10±0.01</td>
<td>0.12±0.02</td>
<td>.31</td>
</tr>
</tbody>
</table>

**ASC score**

<table>
<thead>
<tr>
<th>ASC score</th>
<th>Control</th>
<th>Case</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5 (0.4)</td>
<td>1 (0.3)</td>
<td>.21&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>1</td>
<td>64 (4.9)</td>
<td>16 (4.9)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>241 (18.4)</td>
<td>33 (10.1)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>317 (24.2)</td>
<td>93 (28.4)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>340 (25.9)</td>
<td>102 (31.1)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>199 (15.7)</td>
<td>52 (15.9)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>110 (8.4)</td>
<td>23 (7.0)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>32 (2.4)</td>
<td>6 (1.8)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>4 (0.3)</td>
<td>2 (0.6)</td>
<td></td>
</tr>
</tbody>
</table>

**Coffee, regular drinker (%)**

<table>
<thead>
<tr>
<th>Coffee, regular drinker (%)</th>
<th>Control (%)</th>
<th>Case (%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>791 (60.8)</td>
<td>188 (57.9)</td>
<td>.33</td>
<td></td>
</tr>
</tbody>
</table>

**Caffeine, mg/day**

<table>
<thead>
<tr>
<th>Caffeine, mg/day</th>
<th>Control</th>
<th>Case</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>173.7±3.8</td>
<td>167.3±6.8</td>
<td>.42</td>
<td></td>
</tr>
</tbody>
</table>

**Smoking status**

<table>
<thead>
<tr>
<th>Smoking status</th>
<th>Control</th>
<th>Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never smokers (%)</td>
<td>644 (49.1)</td>
<td>163 (49.7)</td>
</tr>
<tr>
<td>Former smokers (%)</td>
<td>575 (43.8)</td>
<td>149 (45.4)</td>
</tr>
<tr>
<td>Current smokers (%)</td>
<td>93 (7.1)</td>
<td>16 (4.9)</td>
</tr>
</tbody>
</table>

**Duration of regular smoking, years**

<table>
<thead>
<tr>
<th>Duration of regular smoking, years</th>
<th>Control</th>
<th>Case</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.83±0.06</td>
<td>1.67±0.11</td>
<td>.24</td>
<td></td>
</tr>
</tbody>
</table>

**Past oral contraceptive (%)**

<table>
<thead>
<tr>
<th>Past oral contraceptive (%)</th>
<th>Control</th>
<th>Case</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>569 (43.4)</td>
<td>122 (37.2)</td>
<td>.03</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.1. Baseline Characteristics and WHI Study Participation by Control and Case Status (continued)

| Past estrogen-alone therapy (%) | 427 (32.6) | 124 (37.8) | .07 |
| Past estrogen+progestin therapy (%) | 359 (27.4) | 97 (29.6) | .42 |
| **Education, college graduate or above (%)** | 516 (39.7) | 150 (45.7) | .05 |

**Marital status**

| Never married (%) | 49 (3.8) | 5 (1.5) |
| Divorced or separated (%) | 201 (15.4) | 51 (15.6) |
| Widowed (%) | 224 (17.1) | 58 (17.7) |
| Presently married (%) | 799 (61.1) | 211 (64.3) |
| Marriage-like relationship (%) | 34 (2.6) | 3 (0.9) |

**Family income**

| <$10,000-$19,999 (%) | 193 (15.7) | 46 (14.9) |
| $20,000-$49,999 (%) | 549 (44.8) | 125 (40.5) |
| $50,000-$99,999 (%) | 353 (28.8) | 103 (33.3) |
| ≥$100,000 (%) | 131 (10.7) | 35 (11.3) |

**Health insurance (yes) (%)**

| 1245 (95.5) | 315 (96.6) | .36 |

**WHI study part**

| Observational study (OS,%) | 729 (55.6) | 166 (50.6) | .10d |
| Clinical trial (CT,%) | 583 (44.4) | 162 (49.4) |
| **EH trials** | 246 | 68 |
| Control arm (%) | 114 (46.3) | 35 (51.5) | .72f |
| Estrogen alone arm (E-alone, %) | 48 (19.5) | 13 (19.1) |
| Estrogen+progestin arm (E+P, %) | 84 (34.2) | 20 (29.4) |
| **Dietary modification trial (DM)** | 403 | 119 | .60g |

(continued)
Table 4.1. Baseline Characteristics and WHI Study Participation by Control and Case Status (continued)

<table>
<thead>
<tr>
<th></th>
<th>Control arm (%)</th>
<th>Dietary change arm (%)</th>
<th>Calcium and vitamin D trial (CaD)</th>
<th>Calcium carbonate + vitamin D3 arm (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>226 (56.1)</td>
<td>177 (43.9)</td>
<td>326</td>
<td>161 (49.4)</td>
</tr>
<tr>
<td></td>
<td>70 (58.8)</td>
<td>49 (41.2)</td>
<td>81</td>
<td>45 (55.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><img src="image" alt="Table Image" /></td>
<td></td>
</tr>
</tbody>
</table>

Notes: ACS=American Cancer Society; BMI=Body mass index; CaD=Calcium and Vitamin D; CT=Clinical trials; DM=Dietary modification; E+P=Estrogen progestin therapy; E-alone=Estrogen therapy; OS=Observational study. For categorical variables frequency and percentage are provided, for continuous variables, the mean and standard error are provided. P-values are obtained from the likelihood ratio test from conditional logistic regression models for case/control status.

Items marked with a "*" indicate ACS score components.

a BMI category was fit as an ordinal variable
b MET (metabolic equivalent) is the sum of moderate and hard exercise MET hours/week.
c ACS score was also fit as an ordinal variable for the following grouping: 0-2 v 3 v 4 v 5 v 6-8. P-value for the 3 level derived ACS score (Low=0-3, Medium=4,5, High=6-8) is .15.
d Comparison of study part (CT v OS).
e Women could be enrolled on any combination of the EH trials, the DM trial, and the CaD trial, therefore subcategories within the clinical trial part are not mutually exclusive.
f Comparison of EH trial arm (control v E-alone v E+P).
g Comparison of DM trial arm (control v. dietary change).
h Comparison of CaD trial arm (control v Calcium carbonate + vitamin D3).
Table 4.2. Baseline Characteristics by WHI Study Status

<table>
<thead>
<tr>
<th>Variables</th>
<th>Clinical Trials (n=745)</th>
<th>Observational Study (n=895)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>63.0±0.3</td>
<td>62.9±0.2</td>
<td>.70</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>.96</td>
</tr>
<tr>
<td>Black or African American (%)</td>
<td>45 (6.0)</td>
<td>55 (6.2)</td>
<td></td>
</tr>
<tr>
<td>White (%)</td>
<td>674 (90.5)</td>
<td>811 (90.6)</td>
<td></td>
</tr>
<tr>
<td>Other (%)</td>
<td>26 (3.5)</td>
<td>29 (3.2)</td>
<td></td>
</tr>
<tr>
<td>U.S. Region</td>
<td></td>
<td></td>
<td>.64</td>
</tr>
<tr>
<td>Northeast (%)</td>
<td>189 (25.4)</td>
<td>207 (23.1)</td>
<td></td>
</tr>
<tr>
<td>South (%)</td>
<td>177 (23.8)</td>
<td>229 (25.6)</td>
<td></td>
</tr>
<tr>
<td>Midwest (%)</td>
<td>158 (21.2)</td>
<td>200 (22.4)</td>
<td></td>
</tr>
<tr>
<td>West (%)</td>
<td>221 (29.7)</td>
<td>259 (28.9)</td>
<td></td>
</tr>
<tr>
<td>Body mass index (BMI), kg/m²</td>
<td>28.6±0.2</td>
<td>27.4±0.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Body mass index category*</td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Underweight (&lt; 18.5) (%)</td>
<td>4 (0.5)</td>
<td>10 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Normal (18.5 - 24.9) (%)</td>
<td>208 (27.9)</td>
<td>334 (37.3)</td>
<td></td>
</tr>
<tr>
<td>Overweight (25.0 - 29.9) (%)</td>
<td>277 (37.2)</td>
<td>324 (36.2)</td>
<td></td>
</tr>
<tr>
<td>Obesity I (30.0 - 34.9) (%)</td>
<td>153 (20.5)</td>
<td>154 (17.2)</td>
<td></td>
</tr>
<tr>
<td>Obesity II (35.0 - 39.9) (%)</td>
<td>75 (10.1)</td>
<td>41 (4.6)</td>
<td></td>
</tr>
<tr>
<td>Extreme Obesity III (≥40) (%)</td>
<td>28 (3.8)</td>
<td>32 (3.6)</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2. Baseline Characteristics by WHI Study Status (continued)

<table>
<thead>
<tr>
<th>Physical activity* (MET hours/week)</th>
<th>6.29±0.2</th>
<th>7.01±0.56</th>
<th>&lt;.001</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diet</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fruit and vegetables*, servings/day</td>
<td>3.92±0.07</td>
<td>4.22±0.07</td>
<td>.01</td>
</tr>
<tr>
<td>Vegetables, servings/day</td>
<td>2.19±0.05</td>
<td>2.21±0.05</td>
<td>.78</td>
</tr>
<tr>
<td>Fruits, servings/day</td>
<td>1.73±0.04</td>
<td>2.00±0.04</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Total carotenoids*, mcg/day</td>
<td>11,048±220</td>
<td>10,677±199</td>
<td>.21</td>
</tr>
<tr>
<td>Red and processed meat*, servings/day</td>
<td>1.08±0.03</td>
<td>0.82±0.03</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Fish, servings/day</td>
<td>0.26±0.01</td>
<td>0.26±0.01</td>
<td>.83</td>
</tr>
<tr>
<td>Dark fish, servings/day</td>
<td>0.03±0.00</td>
<td>0.04±0.00</td>
<td>.03</td>
</tr>
<tr>
<td>White fish, servings/day</td>
<td>0.06±0.00</td>
<td>0.07±0.00</td>
<td>.02</td>
</tr>
<tr>
<td>Soy, servings/day</td>
<td>0.02±0.00</td>
<td>0.03±0.01</td>
<td>.06</td>
</tr>
<tr>
<td>Dairy, servings/day</td>
<td>1.92±0.05</td>
<td>1.83±0.04</td>
<td>.18</td>
</tr>
<tr>
<td>Whole grains, servings/day</td>
<td>1.16±0.03</td>
<td>1.22±0.03</td>
<td>.11</td>
</tr>
<tr>
<td>Proportion of grains as whole grains*</td>
<td>0.25±0.00</td>
<td>0.28±0.01</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Fats/oils, total g/day</td>
<td>71.86±1.30</td>
<td>53.93±1.17</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Saturated fats, g/day</td>
<td>24.46±0.48</td>
<td>18.00±0.42</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Monounsaturated fats, g/day</td>
<td>27.26±0.50</td>
<td>20.44±0.46</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Polyunsaturated fats, g/day</td>
<td>14.55±0.28</td>
<td>11.17±0.24</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Trans fats, g/day</td>
<td>5.23±0.12</td>
<td>3.77±0.10</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Alcohol*, drinks/day</td>
<td>0.33±0.02</td>
<td>0.37±0.02</td>
<td>.27</td>
</tr>
<tr>
<td>Beer, drinks/day</td>
<td>0.04±0.01</td>
<td>0.04±0.01</td>
<td>.62</td>
</tr>
<tr>
<td>Wine, drinks/day</td>
<td>0.19±0.02</td>
<td>0.22±0.02</td>
<td>.14</td>
</tr>
</tbody>
</table>

(continued)
Table 4.2. Baseline Characteristics by WHI Study Status (continued)

<table>
<thead>
<tr>
<th>Liquor, drinks/day</th>
<th>0.10±0.01</th>
<th>0.10±0.01</th>
<th>.90</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1 (0.1)</td>
<td>5 (0.6)</td>
<td>.</td>
</tr>
<tr>
<td>1</td>
<td>53 (7.1)</td>
<td>27 (3.0)</td>
<td>.</td>
</tr>
<tr>
<td>2</td>
<td>145 (19.5)</td>
<td>129 (14.4)</td>
<td>.</td>
</tr>
<tr>
<td>3</td>
<td>204 (27.4)</td>
<td>206 (23.0)</td>
<td>.</td>
</tr>
<tr>
<td>4</td>
<td>190 (25.5)</td>
<td>252 (28.2)</td>
<td>.</td>
</tr>
<tr>
<td>5</td>
<td>95 (12.8)</td>
<td>156 (17.4)</td>
<td>.</td>
</tr>
<tr>
<td>6</td>
<td>45 (6.0)</td>
<td>88 (9.8)</td>
<td>.</td>
</tr>
<tr>
<td>7</td>
<td>10 (1.3)</td>
<td>28 (3.1)</td>
<td>.</td>
</tr>
<tr>
<td>8</td>
<td>2 (0.3)</td>
<td>4 (0.5)</td>
<td>.</td>
</tr>
<tr>
<td><strong>Coffee, regular drinker (%)</strong></td>
<td>460 (62.3)</td>
<td>519 (58.4)</td>
<td>.11</td>
</tr>
<tr>
<td><strong>Caffeine, mg/day</strong></td>
<td>183.3±5.3</td>
<td>163.3±4.2</td>
<td>.01</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
<td></td>
<td>.15</td>
</tr>
<tr>
<td>Never smokers (%)</td>
<td>381 (51.1)</td>
<td>426 (47.6)</td>
<td></td>
</tr>
<tr>
<td>Former smokers (%)</td>
<td>310 (41.6)</td>
<td>414 (46.3)</td>
<td></td>
</tr>
<tr>
<td>Current smokers (%)</td>
<td>54 (7.3)</td>
<td>55 (6.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Duration of regular smoking, years</strong></td>
<td>1.75±0.08</td>
<td>1.84±0.07</td>
<td>.42</td>
</tr>
<tr>
<td>Past oral contraceptive (%)</td>
<td>310 (41.6)</td>
<td>381 (42.6)</td>
<td>.70</td>
</tr>
<tr>
<td>Past estrogen-alone therapy (%)</td>
<td>241 (32.4)</td>
<td>310 (34.6)</td>
<td>.33</td>
</tr>
<tr>
<td>Past estrogen+progestin therapy (%)</td>
<td>172 (23.1)</td>
<td>284 (31.7)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Education, college graduate or above (%)</td>
<td>276 (37.3)</td>
<td>390 (43.9)</td>
<td>.01</td>
</tr>
</tbody>
</table>
### Table 4.2. Baseline Characteristics by WHI Study Status (continued)

<table>
<thead>
<tr>
<th>Marital status</th>
<th>.73</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never married (%)</td>
<td>25 (3.4)</td>
</tr>
<tr>
<td>Divorced or separated (%)</td>
<td>111 (14.9)</td>
</tr>
<tr>
<td>Widowed (%)</td>
<td>119 (16.0)</td>
</tr>
<tr>
<td>Presently married (%)</td>
<td>470 (63.3)</td>
</tr>
<tr>
<td>Marriage-like relationship (%)</td>
<td>18 (2.4)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Family income</th>
<th>.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$10,000-$19,999 (%)</td>
<td>113 (15.7)</td>
</tr>
<tr>
<td>$20,000-$49,999 (%)</td>
<td>325 (45.2)</td>
</tr>
<tr>
<td>$50,000-$99,999 (%)</td>
<td>209 (29.1)</td>
</tr>
<tr>
<td>$\geq 100,000 (%)</td>
<td>72 (10.0)</td>
</tr>
</tbody>
</table>

| Health insurance (yes) (%)     | 704 (95.0) | 856 (96.3) | .20  |

Notes: ACS=American Cancer Society; BMI=Body mass index; CT=Clinical trials; OS=Observational study. For categorical variables frequency and percentage are provided and the p-value is obtained from a chi-square test to compare the CT and OS groups. For continuous variables, the mean and standard error are provided and the p-value is obtained from a two-sided t-test to compare women on CT and OS. Items marked with a “*” indicate ACS score components.

* BMI category was fit as an ordinal variable.

†MET (metabolic equivalent) is the sum of moderate and hard exercise MET hours/week.

‡The p-value corresponds to the ACS score being fit as an ordinal variable for the following grouping:0-2 v 3 v 4 v 5 v 6-8.

§Comparison of study part (CT v OS).
### Table 4.3 Conditional Logistic Regression Multivariable Modeling for CLL/SLL

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular coffee drinking (yes v. no)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On CT</td>
<td>0.730</td>
<td>0.509-1.046</td>
<td>.08</td>
</tr>
<tr>
<td>On OS</td>
<td>0.997</td>
<td>0.701-1.420</td>
<td>.78</td>
</tr>
<tr>
<td>Any past estrogen therapy use (yes v. no)</td>
<td>1.316</td>
<td>1.015-1.706</td>
<td>.04</td>
</tr>
<tr>
<td>Any past oral contraceptive use (yes v. no)</td>
<td>0.735</td>
<td>0.561-0.963</td>
<td>.03</td>
</tr>
<tr>
<td>Obesity status (≥30 BMI vs. &lt;30 BMI)</td>
<td>0.735</td>
<td>0.561-0.963</td>
<td>.02</td>
</tr>
<tr>
<td>Ever smoker (yes v. no)</td>
<td>1.006</td>
<td>0.785-1.299</td>
<td>.96</td>
</tr>
<tr>
<td>U.S. Region b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>1.231</td>
<td>0.870-1.740</td>
<td>.60</td>
</tr>
<tr>
<td>South</td>
<td>1.195</td>
<td>0.847-1.686</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>1.077</td>
<td>0.750-1.547</td>
<td></td>
</tr>
</tbody>
</table>

Notes: BMI=Body mass index; CT=Clinical trials; OS=Observational study. Conditional logistic regression was used to obtain the odds ratios and corresponding 95% confidence intervals for CLL/SLL. 

*P*-value corresponds to the 2X2 interaction term of any coffee use (yes, vs. no) by study part (CT,OS) in the model.

bWest is reference group.
Chapter 5: Hormone Exposure and Incidence of Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma

5.1 Background

Although leukemia is a relatively rare form of cancer in adults (approximately 3% of all cancers), there are about 240,000 new cases diagnosed annually worldwide (43,800 in the U.S.) and approximately 200,000 deaths associated with leukemia (23,300 in the U.S.) each year.\(^7\)\(^8\) Leukemia is ranked fifth in person-years of life lost due to cancer, directly behind breast and pancreatic cancer.\(^9\) According to the National Cancer Institute’s Surveillance, Epidemiology and End Results (SEER) Program, during recent decades (between 1975 and 2010), there has been a trend of increasing incidence of leukemia among women— with no clear explanation to support this trend.

In industrialized countries, Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma (CLL/SLL) is the most common type of adult leukemia, one that is associated with aging, with causes largely unknown. Despite being a prevalent type of leukemia, CLL/SLL is a relatively rare form of cancer (approximately 1% of all cancers). There are still about 15,720 new CLL/SLL cases diagnosed each year in the U.S – mainly in older adults (the lifetime risk of CLL/SLL is 0.52%, median age at diagnosis: 70 years).\(^7\)\(^8\)\(^10\) Patients at this age have been subjected to several decades of DNA
damage from endogenous and exogenous sources, which emphasizes the need to recognize the multiple types of exposures that may contribute to CLL/SLL risk. Since CLL/SLL is twice as common among males, studying women may reveal potentially protective factors relative to female hormone exposures, either intrinsic (such as a result of higher estrogen levels due to obesity) or extrinsic, by oral contraceptive (OC) or hormone therapy (HT).

There are hypotheses for various biological mechanisms by which estrogen could be associated with leukemia. While there are few epidemiologic studies that have explored the potential effects of estrogen use on CLL/SLL incidence, it has been established that estrogen receptors (ERs) are present on the lymphocytes of CLL/SLL patients. The biological significance of this finding is not well understood. Although not recognized generally as a hormone-controlled cancer, the most recent research on Non-Hodgkin Lymphoma (NHL; which includes CLL/SLL as subtype) by Yakimchuck et al. supports that it is. The authors showed that 1.) the activation of ERβ (the dominant ER expressed in lymphoid cells) in malignant B-cell lymphoid cells acts as a tumor suppressor; 2.) this growth suppressing effect is specific to the tumor cells and not the microenvironment and; 3.) that tumor suppression is achieved by ERβ signaling which inhibits angiogenesis and lymphangiogenesis.

Studies investigating HT use and leukemia have been lacking. Cerhan et al. conducted a study investigating past and current HT use and NHL and CLL/SLL. This study only had 63 CLL/SLL cases and did not find any relationship between HT use and CLL/SLL. However, the authors did identify that current HT use was positively
associated with the incidence of NHL among the 258 NHL cases studied. A recently
published study by Kato et. al analyzed women from the Women's Health Initiative
(WHI) HT randomized studies and did not identify an association between randomized
treatment groups versus placebo with NHL incidence.\textsuperscript{194} As HT, particularly treatment
with estrogen-alone (E-alone), is commonly prescribed for women who undergo
oophorectomy, an animal study by Kalu et al. showed that oophorectomy was associated
with an increase in the number of lymphoid cells in the BM.\textsuperscript{195}

With respect to OC use, only a few studies have investigated risk of leukemia. It
has been known for some time that, during pregnancy, sex steroids (estrogen and
progesterone) are responsible for the reduction in B-lymphopoiesis and this finding was
further supported by evaluating exogenous estrogen therapy in animal models.\textsuperscript{196} With
respect to the estrogenic effects of OCs on lymphocytes, there have been a few \textit{in vitro}
studies that demonstrated a basis for potential biological mechanisms. Pozsonyi et al.
found that upon treating healthy women's cells with a mitogen \textit{in vitro}, those who took
OC had a decrease in lymphoblastic transformation, compared to women who did not,
suggesting an inhibitory effect of estrogen on interleukin-2 production.\textsuperscript{197} Another study
showed that OCs influence lymphocytes, in particular cytotoxic lymphocytes and B cells.
\textsuperscript{198} Specifically among OC users, the levels of CD3+ and CD8+ cells were higher,
suggesting either effector or suppressor (or both) activity, whereas the number of NK
cells decreased. A recent epidemiologic study by Lu et al. found a trend for an inverse
association between OC use and the incidence of NHL. However, this study only
included 110 CLL/SLL cases and the CLL/SLL subset analysis was likely underpowered,
and thus, no association between OC use and CLL/SLL was revealed. 199

We conducted a focused investigation of the association between EH use (HT and OC) and CLL/SLL using data from the WHI, a large prospective study of post-menopausal women. We evaluated the association between EH exposure and the risk of CLL/SLL. We also investigated whether there is a differential effect according to timing (past versus current) of EH exposure and the type of exposure (OC, E-alone, estrogen plus progestin (E+P)).

5.2 Methods

Study Design

The WHI was designed to address the major causes of morbidity and mortality in postmenopausal women and includes four clinical trials (CTs) and an observational study (OS). 176 Details of the scientific rationale, eligibility requirements, and baseline participant characteristics of the WHI have been published elsewhere. 177,178,179,180,181 Briefly, a total of 161,808 women, 50–79 years of age, were recruited at 40 clinical centers throughout the United States between September 1, 1993 and December 31, 1998. The four CTs were: two HT trials (27,347 women), a dietary modification trial (DM; 48,835 women), and a calcium/vitamin D supplementation trial (CaD; 36,282 women).

With respect to the two HT trials, one study investigated E-alone in post-menopausal women without a uterus with the experimental arm taking a daily dose of estrogen in the form of conjugated equine estrogen (CEE) for an average of about 6 years. The other study looked at E+P in post-menopausal women who still had their uterus; women on the
experimental arm of this study took a daily dose of CEE plus a progestin (medroxyprogesterone acetate) for an average of about 5 years. Both HT trials had placebo control arms. Participants in the OS included 93,676 women who were screened for the clinical trials but proved to be ineligible or unwilling to participate or who were recruited through a direct invitation for the OS. A large percentage (72.8%) of women in the OS took HT (E-alone or E+P) while enrolled in the study. All WHI participants signed informed consent and were followed prospectively. The WHI study was overseen by institutional review boards at all 40 clinical centers and at the coordinating center, and by a study-wide data-safety monitoring board.

This study was a 1:4 nested case-control study within the CT and the OS of the WHI. The following participants were excluded from the original cohort of 161,808 for this analysis: 24,654 women who had a history of cancer at baseline, 7148 women who had missing main exposure variables and 405 women with a new leukemia diagnosis (other than CLL/SLL) during the study. This resulted in a sample size of 129,601 subjects, of whom 328 were diagnosed with CLL/SLL during the WHI follow-up. Applying a random selection of four controls for each case, matched by age (5-year window) and race, provided 1312 controls and 1640 subjects for statistical analyses.

**Measures**

*Hormone therapy exposure measurement*

Current users of HT were defined as women who reported using HT at baseline in the OS (confirmed at the 3-year follow-up visit), or women (not participating on the HT trials) using HT at baseline in the DM or CaD trials (confirmed at 3-years) or women
assigned to HT use in the HT trials. Both E-alone and E+P use were included in the evaluation of current HT use. With the exception of HT CT participants (HT CT women with intact uterus received conjugated equine estrogens (CEE; 0.625 mg/d) plus medroxyprogesterone acetate (MPA; 2.5 mg/d) or placebo and those with prior hysterectomy received CEE alone (0.625 mg/d) or placebo), all current HT use data were collected via self-administered questionnaire of current medication use.

Data on past use of HT were collected as part of a self-administered questionnaire at baseline for all participants (OS and CT) in the study. Past users of HT were defined as women who had used HT in the past, irrespective of their current use of HT on the study. Therefore, women could be past, current, past and current users, or never users of HT. Both E-alone and E+P use were included in the evaluation of past HT use. Data on age at first use and duration of use were also collected.

*Oral contraceptive exposure measurement*

Data on past use of OC were collected at baseline for all participants (OS and CT) in the study. Past users of OC were defined as women who had used OC in the past for any duration of time (i.e., “ever use of OC”). Specifically, women were asked in a self-administered questionnaire if they had ever used OC. Data on age at first use and duration of use were also collected.

*Follow-up and ascertainment of cases*

According to the World Health Organization (WHO), since 2008, small lymphocytic lymphoma (SLL) and CLL/SLL have been considered one disease and one
entity for disease classification, as malignant cells in both diagnoses exhibit the same immunophenotype. In our study, we followed the most current guidelines (WHO 2008) that include SLL cases as CLL/SLL, and therefore all references to CLL/SLL within our study include both CLL/SLL and SLL cases.

Incident CLL/SLL cases were identified by self-administered questionnaires (administered annually in the WHI CT after 2005, and annually in the WHI OS throughout the study), with all cases confirmed by medical record review. All CLL/SLL cases then were coded centrally in accordance with the Surveillance Epidemiology and End Results (SEER) coding guidelines. For identification of cases, participants were followed up to CLL/SLL diagnosis, date of death, loss to follow-up, or end of WHI CT or OS follow-up, whichever occurred first.

Statistical Analysis

Baseline characteristics were summarized for cases and controls displaying mean and standard error for continuous variables and frequencies with percentages for categorical variables. To compare the groups for variables not used in the matching, univariable conditional logistic regression models were generated and the likelihood-ratio p-values were reported. Current HT use, obtained at the time of randomization (for HT trial participants) and at the three-year follow-up (for all other women not part of the HT trial), were summarized.

Univariable conditional logistic regression models estimated CLL/SLL risk and past OC use, past HT use, current HT use and other important risk factors. To evaluate
important hormone use related risk factors for CLL/SLL incidence in the presence of other important risk factors, multivariable conditional logistic regression analysis was conducted. Odds ratios (ORs) and corresponding 95% confidence intervals and p-values were calculated for the matched-pair cohort data and ORs were used to estimate risk ratios. Interactions between EH use and other risk factors (i.e., obesity, smoking status) were evaluated by including a multiplicative interaction term in the regression model to evaluate any effect modification. Analyses were performed using SAS 9.3 (SAS Institute, Cary NC).

5.3 Results

Descriptive Data Analysis

A total of 328 CLL/SLL cases meeting study inclusion criteria (i.e., no previous history of cancer, non-missing main exposure variables) were identified, during a mean follow-up period of 12.7 years. Baseline characteristics according to case - control status are shown in Table 5.1. Since age and race were both used in matching controls to cases, these variables were, not surprisingly, evenly distributed between the groups by definition. With respect to other important baseline characteristics, body mass index (BMI) measured at the baseline was lower for cases (P=.03). Defining obesity status as BMI ≥ 30, it was observed that cases were significantly less likely to be obese (P=.03). Cases also showed a trend for higher bilateral oophorectomy (BOO) history (21.7% v 17.4%; P=.07), and to be younger at menopause (P=.08), compared to controls. In addition, cases were more often college graduates compared to controls (P=.05). Smoking status (i.e., “Smoked at least 100 cigarettes ever”), region of the U.S. they resided in, and
all other demographic factors were similar between cases and controls. Data on EH exposure according to case-control status are presented in Table 5.2. When any past HT was considered (either E-alone or E+P, or both), a higher percentage of cases took HT (60.4% v. 55.3%; P=.09) compared to controls, suggesting an adverse effect on CLL/SLL. Age at first HT use occurred by about a year sooner among cases than controls (P=.09) and duration of HT use tended to be about a year longer (P=.13). Further examining the type of hormone, past E-alone use showed a trend to be more frequent among cases (37.8% v. 32.6%; P=.07) whereas E+P was similar between cases and controls (29.6% v. 27.4%; P=.42). The difference between age at first HT use was driven by cases taking E-alone therapy a year earlier than controls (P=.13) and similarly the duration of E-alone therapy was nearly 16 months longer among cases than controls (P=.15), where as there were no differences detected for these measures among E+P therapy users (age at first E+P: P=.46, duration of past E+P: P=.68). With respect to OC, past use was significantly lower among cases, indicating a protective effect of OC use against CLL/SLL (P=.04). Age at first use of OC (P=.75) and duration of OC use (P=.45) were similar between cases and controls. Current HT use—irrespective of type—was not associated with CLL/SLL (P=.61) and this finding was consistent with the lack of association found when considering type of therapy (E-alone: P=.21; E+P: P=.57).

**Hormonal Exposures and the CLL/SLL Risk**

To further describe the relationship between EH use and risk of CLL/SLL, we estimated the OR for CLL/SLL risk along with its 95% confidence interval. We arrived at three final multivariable models, each showing the impact of past EH through additional
level of detail (Table 5.3). Model 1 evaluated any past HT use as its main variable and showed that any past HT use resulted in a 32% increase in the risk of developing CLL/SLL (OR=1.32, 95% CI: 1.02-1.70, P=.04), past OC use and being obese, however, were associated with significant reductions in the risk for CLL/SLL (Past OC: OR= 0.73, 95% CI: 0.56-0.96, P=.02; Obesity: OR= 0.73, 95% CI: 0.55-0.97, P=.03), after adjusting for smoking status, region and study arm (Figure 5.1). The interaction term of obesity status and OC use attained marginal significance (P=.10) and through stratified analyses revealed that obese women who took OC in the past had about half the risk of CLL/SLL compared to women who did not take OC or took OC but were not obese. No other significant effect modification was found between HT use and other variables for the risk of CLL/SLL. Study arm was statistically significant in the final multivariable model (In Models 1 and 2: P=.05) for any past HT use indicating that women enrolled in the CT had a 29% increase in the risk of developing CLL/SLL, compared to OS participants. Upon additional analysis, there was no effect modification of study part (CT vs. OS) and CLL/SLL incidence with any of the other variables. Yet, we noted that a higher percentage of women on the CT were obese (34%) compared to the OS (25%), therefore, in the final models which also included obesity as a risk factor, controlling for this slight disproportion of obese women across study parts became critical in order to obtain a reliable estimate of CLL/SLL risk purely due to being obese.

In Model 2 we evaluated the specific type of past HT (E-alone v. E+P) in order to distill any differential effect. This model suggested that the adverse impact of past HT use found in Model 1 was mainly driven by past E-alone use (OR= 1.368, 95% CI: 1.05-
1.78, P= .02). In Model 3, we investigated past E-alone use relative to BOO (for which E-alone is the most commonly prescribed therapy). We found that any use of past E-alone, irrespective of BOO was associated with an increased risk of CLL/SLL. Of note, neither smoking status, nor region were significant in the three final models, however they were kept in the final models to control for potential confounding.

5.4 Discussion

In this nested case-control study, designed using a large prospective cohort of postmenopausal women, we found that women with past, but not current HT exposure, had a significantly elevated risk of developing CLL/SLL. When further investigated, this association was mainly due to E-alone, whereas E+P did not have a statistically significant effect on the risk of CLL/SLL. The risk of CLL/SLL increased by about 30% among women with past E-alone use. In our study, 44% of women who took E-alone in the past were likely prescribed this therapy due to having had a hysterectomy with a common companion procedure, BOO. As most estrogen is produced in the ovaries, when both ovaries are removed through BOO, estrogen levels fall sharply, so past E-alone users were likely short on estrogen. Estrogen therapy was given to replace lowered estrogen levels and likely did not meet normal levels despite therapy. Therefore, women taking E-alone still likely had suboptimal estrogen levels compared to other women. This would support our finding with respect to women who took E-alone in the past being at higher risk for CLL/SLL, given that we found that OCs were protective against CLL/SLL most likely due to their high estrogen levels. We showed that past E-alone use, irrespective of BOO, resulted in increased rates of CLL/SLL. On the other hand, women
taking E+P were mainly ones with intact uterus and ovaries (only 8% of past E+P users went through BOO), and were most likely prescribed E+P to relieve menopausal symptoms such as hot flashes. Therefore, past E+P users were more likely to still have had adequate levels of intrinsic estrogen especially after HT which explains why we did not see a statistically significant association with E+P and CLL/SLL. These findings are contrary to those from a previous study by Lu et. al, who found that women who had a BOO and took E-alone or E+P had a lower risk of B-cell NHL. One explanation for this discrepancy could be that the Lu et. al study only included 93 CLL/SLL cases for their HT analysis and their study did not separately focus on CLL/SLL relative to BOO and past HT, but only considered the larger group of B-cell NHL, diluting any potential findings specific to CLL/SLL.

We also observed a protective effect for CLL/SLL with past OC use: women who used OC in the past had a 27% reduction of CLL/SLL compared to women who did not use OC. In our study- similarly to current trends in industrialized populations- past OC use was common, with 42% of women being exposed, indicating this finding could have a large impact due to the wide prevalence of this exposure. In addition, during the period the cohort of women who later enrolled in the WHI study were taking OC, the dose of estrogen (as well as progestin) in the pills was significantly higher than currently prescribed OC drugs. Typical estrogen doses in OCs prescribed in the 1960s-1980s ranged from 30 mg-100 mg, while current estrogen doses in OCs range from 30 mg-50 mg. Oral contraceptives are not only prescribed to prevent unwanted pregnancy, but also for non-contraceptive benefits, such as to treat menstrual irregularities, peri-
menopausal vasomotor symptoms, acne, and hirsutism. In light of our findings for CLL/SLL and previous findings relative to OCs protective effect against ovarian and endometrial cancers, the use of OCs could also be evaluated for chemoprevention of these cancers. Studies evaluating the net benefits versus potential adverse effects (cardiovascular, breast, and cervical cancer) will need to be conducted, by also including CLL/SLL as an outcome.

With respect to the biological findings of estrogen receptors on lymphocytes (B and T) in CLL/SLL, it has recently been shown that the majority of patients with CLL/SLL express ERβ1 or ERβ2, suggesting that estrogen receptors may be important players in the development of CLL/SLL and also these ERs may be potential targets for cancer therapy. Relative to normal BM cells, in leukemia there is a repressed expression of the Estrogen Receptor 1 gene (ESR1), located on chromosome 6q25. Additionally, the expression of ESR1 is regulated by the p53 tumor suppressor protein and in breast cancer cells it has been found that p53 binds to both methylated and unmethylated CpG islands of ESR1. With respect to epigenetics particular to leukemia, in about one half of chronic leukemia cases there is DNA hypermethylation on the ESR1 CpG islands, compared to normal non-malignant cells. These findings suggest that hypermethylation of the ESR1 CpG islands correlates with ESR1 silencing, similarly to what has been established relative to other genes in AML, where hypermethylated genes include tumor suppressors. In addition to ESR1, there are other genes with prevalent promoter DNA hypermethylation in newly diagnosed CLL/SLL patients. For example, the SFRP-1 gene is hypermethylated in CLL/SLL and is known to be estrogen-
inducible. Also, the \textit{RASSF1A} gene promoter methylation levels have been found to positively correlate with estrogen receptor expression in breast cancer patients, and this gene is frequently hypermethylated in newly diagnosed CLL/SLL patients. A recent study also found that in women who took OCs, global methylation levels in their white blood cells were lower compared to women who did not take OCs. In addition to ER, progesterone receptor (PR) activity has also been observed on B-lymphocytes of CLL/SLL patients, though in lesser extent, in about a quarter of cases, however the impact of PRs relative to the development of CLL/SLL in relation to EH exposures has not been studied.

We did not find an association specific to current HT use and CLL/SLL incidence. The most likely explanation for this is that the effect of current HT was not observed in our study since CLL/SLL is most likely to develop as a result of longer term exposures. However, it is important to note that the majority of past HT users were also current users, so it is difficult to separate the treatment effect of HT by time period in our study. Particularly, 62% of past E-alone users went on to also become current E-alone users.

Our research has multiple strengths including 1) the nested case-control design that enabled us to efficiently assess the risk of CLL/SLL which is a rare cancer and is therefore difficult to study in a cohort setting, 2) its focus on women relative to hormone exposure and association with CLL/SLL in an adequately powered study for the first time by including both the CT and OS parts of WHI, 3) confirmed CLL/SLL diagnosis, and 4) use of a well-designed questionnaire with detail on past EH use. However, there are some
possible limitations to our study, such as past OC and HT use- our main exposure variables- were self-reported and some level of misclassification was possible. Also, we did not have data available on ERs in the leukemic blasts of the cases, which could have been supportive of some of our findings. Finally, we only had past weight history reported on a subset of the women (OS participants) in our study, therefore we could not analyze concurrent obesity status relative to OC use, only more recent obesity status measured at study baseline. However, upon a subset analysis of OS participants with available past weight data, we found that weight at age 35 years, at approximately the tail end of OC use, significantly correlated with weight at study baseline ($r=0.7$, $P<.001$), which confirms that obesity status measured at study baseline functions as an appropriate surrogate in our study.

In conclusion, postmenopausal women with past HT use (mainly due to E-alone use) had an increased in CLL/SLL incidence by about 32%. On the other hand, we found past OC use and obesity to be protective for CLL/SLL. Past OC use resulted in about a 27% reduction in the risk of getting CLL/SLL. In addition, women who reported past OC use and were obese had about half the rate of getting CLL/SLL when compared to women who did not take OC or took OC but were not obese. Current HT use, irrespective of type (E-alone or E+P) was not associated with the incidence of CLL/SLL. Biological studies are needed to support our findings, preferably more in line with current medical practice where lower doses of estrogen in OCs are common practice as is less frequent HT use, compared to what women in the WHI cohort were exposed to. Finally, from our study it unclear whether EH use is a contributor for the increasing trend in CLL/SLL.
incidence among women and further research will need to be conducted to comprehensively evaluate all conceivable risk factors for CLL/SLL in women.
Table 5.1 Baseline Characteristics and WHI Study Participation by Control and Case Status

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=1312)</th>
<th>Cases (n=328)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>62.9±0.2</td>
<td>63.2±0.4</td>
<td>Matched</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American (%)</td>
<td>80 (6.1)</td>
<td>20 (6.1)</td>
<td></td>
</tr>
<tr>
<td>White (%)</td>
<td>1188 (90.6)</td>
<td>297 (90.6)</td>
<td></td>
</tr>
<tr>
<td>Other (%)</td>
<td>44 (3.3)</td>
<td>11 (3.3)</td>
<td></td>
</tr>
<tr>
<td>U.S. Region</td>
<td></td>
<td></td>
<td>.55</td>
</tr>
<tr>
<td>Northeast (%)</td>
<td>310 (23.6)</td>
<td>86 (26.2)</td>
<td></td>
</tr>
<tr>
<td>South (%)</td>
<td>320 (24.4)</td>
<td>86 (26.2)</td>
<td></td>
</tr>
<tr>
<td>Midwest (%)</td>
<td>291 (22.2)</td>
<td>67 (20.4)</td>
<td></td>
</tr>
<tr>
<td>West (%)</td>
<td>391 (29.8)</td>
<td>89 (27.1)</td>
<td></td>
</tr>
<tr>
<td>Body mass index (BMI), kg/m²</td>
<td>28.1±0.2</td>
<td>27.3±0.4</td>
<td>.03</td>
</tr>
<tr>
<td>Obese (% BMI ≥30)</td>
<td>402 (30.6)</td>
<td>81 (24.7)</td>
<td>.03</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td>.84</td>
</tr>
<tr>
<td>Never smokers (%)</td>
<td>644 (49.1)</td>
<td>163 (49.7)</td>
<td>.</td>
</tr>
<tr>
<td>Former smokers (%)</td>
<td>575 (43.8)</td>
<td>149 (45.4)</td>
<td></td>
</tr>
<tr>
<td>Current smokers (%)</td>
<td>93 (7.1)</td>
<td>16 (4.9)</td>
<td></td>
</tr>
</tbody>
</table>

128 (continued)
Table 5.1 Baseline Characteristics and WHI Study Participation by Control and Case Status (continued)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=1312)</th>
<th>Cases (n=328)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education, college graduate or above (%)</strong></td>
<td>516 (39.7)</td>
<td>150 (45.7)</td>
<td>.05</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married (%)</td>
<td>49 (3.8)</td>
<td>5 (1.5)</td>
<td>.10</td>
</tr>
<tr>
<td>Divorced or separated (%)</td>
<td>201 (15.4)</td>
<td>51 (15.6)</td>
<td></td>
</tr>
<tr>
<td>Widowed (%)</td>
<td>224 (17.1)</td>
<td>58 (17.7)</td>
<td></td>
</tr>
<tr>
<td>Presently married (%)</td>
<td>799 (61.1)</td>
<td>211 (64.3)</td>
<td></td>
</tr>
<tr>
<td>Marriage-like relationship (%)</td>
<td>34 (2.6)</td>
<td>3 (0.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Family income</strong></td>
<td></td>
<td></td>
<td>.40</td>
</tr>
<tr>
<td>&lt;$10,000-$19,999 (%)</td>
<td>193 (15.7)</td>
<td>46 (14.9)</td>
<td></td>
</tr>
<tr>
<td>$20,000-$49,999 (%)</td>
<td>549 (44.8)</td>
<td>125 (40.5)</td>
<td></td>
</tr>
<tr>
<td>$50,000-$99,999 (%)</td>
<td>353 (28.8)</td>
<td>103 (33.3)</td>
<td></td>
</tr>
<tr>
<td>≥$100,000 (%)</td>
<td>131 (10.7)</td>
<td>35 (11.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Health insurance (yes) (%)</strong></td>
<td>1245 (95.5)</td>
<td>315 (96.6)</td>
<td>.36</td>
</tr>
<tr>
<td><strong>Number of full-term pregnancies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never pregnant</td>
<td>110 (8.4)</td>
<td>25 (7.7)</td>
<td>.52</td>
</tr>
<tr>
<td>Pregnant, but not full-term</td>
<td>31 (2.4)</td>
<td>3 (0.9)</td>
<td></td>
</tr>
<tr>
<td>1 full-term pregnancy</td>
<td>115 (8.8)</td>
<td>27 (8.3)</td>
<td></td>
</tr>
<tr>
<td>2 full-term pregnancies</td>
<td>337 (25.8)</td>
<td>88 (26.9)</td>
<td></td>
</tr>
</tbody>
</table>

129 (continued)
### Table 5.1 Baseline Characteristics and WHI Study Participation by Control and Case Status (continued)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=1312)</th>
<th>Cases (n=328)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥3 full-term pregnancies</td>
<td>712 (54.6)</td>
<td>184 (56.3)</td>
<td></td>
</tr>
<tr>
<td>Age at menopause, years</td>
<td>48.5±0.2</td>
<td>47.8±0.4</td>
<td>.08</td>
</tr>
<tr>
<td>Bilateral oophorectomy (BOO, yes) (%)</td>
<td>223 (17.4)</td>
<td>70 (21.4)</td>
<td>.07</td>
</tr>
</tbody>
</table>

**WHI study part**

Observational study (OS,%)  
729 (55.6)  
166 (50.6)  
.10<br><br>
Clinical trial<sup>b</sup> (CT,%)  
583 (44.4)  
162 (49.4)  

**EH trials**  
Control arm (%)  
114 (46.3)  
35 (51.5)  
.72<br><br>
Estrogen alone arm (E-alone, %)  
48 (19.5)  
13 (19.1)  

Estrogen+progestin arm (E+P, %)  
84 (34.2)  
20 (29.4)  

Dietary modification trial (DM)  
Control arm (%)  
226 (56.1)  
70 (58.8)  
.60<br><br>
Dietary change arm (%)  
177 (43.9)  
49 (41.2)  

Calcium and vitamin D trial (CaD)  
Control arm (%)  
165 (50.6)  
36 (44.4)  
.32<br><br>
Calcium carbonate + vitamin D3 arm  
161 (49.4)  
45 (55.6)  

Notes: BMI=Body mass index; BOO=Bilateral oophorectomy; CaD=Calcium and Vitamin D; CT=Clinical trials; DM=Dietary modification; E+P=Estrogen progestin therapy; E-alone=Estrogen therapy; OC=Oral contraceptive; OS=Observational study. For categorical variables frequency and percentage are provided and for continuous variables, the mean and standard error are provided. P-values are obtained from a conditional logistic regression model to compare the two groups.

<sup>a</sup> 2x2 comparison of study part (CT v OS) by CLL/SLL status (case v controls).

<sup>b</sup> Women could be enrolled on any combination of the EH trials, the DM trial, and the CaD trial, therefore subcategories within the clinical trial part are not mutually exclusive.

<sup>c</sup> 3x2 comparison of EH trial arm (control v E-alone v E+P) by CLL/SLL status (case v controls).

<sup>d</sup> 2x2 comparison of DM trial arm (control v. dietary change) by CLL/SLL status (case v controls).

<sup>e</sup> 2x2 comparison of CaD trial arm (control v Calcium carbonate + vitamin D3) by CLL/SLL status (case v controls).
Table 5.2 Exogenous Hormone Use by Control and Case Status

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=1312)</th>
<th>Cases (n=328)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Past hormone therapy (HT) use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any past HT&lt;sup&gt;ab&lt;/sup&gt; (%)</td>
<td>725 (55.3)</td>
<td>198 (60.4)</td>
<td>.09</td>
</tr>
<tr>
<td>Age at first HT (years)</td>
<td>50.4±0.29</td>
<td>49.3±0.56</td>
<td>.09</td>
</tr>
<tr>
<td>Duration of past HT (years)</td>
<td>8.6±0.29</td>
<td>9.6±0.62</td>
<td>.13</td>
</tr>
<tr>
<td>Any past estrogen alone (E-alone) therapy (%)</td>
<td>427 (32.6)</td>
<td>124 (37.8)</td>
<td>.07</td>
</tr>
<tr>
<td>Age at first E-alone therapy (years)</td>
<td>48.6±0.39</td>
<td>47.4±0.71</td>
<td>.13</td>
</tr>
<tr>
<td>Duration of past E-alone therapy (years)</td>
<td>9.8±0.42</td>
<td>11.1±0.85</td>
<td>.15</td>
</tr>
<tr>
<td>Any past estrogen plus progestin (E+P) therapy (%)</td>
<td>359 (27.4)</td>
<td>97 (29.6)</td>
<td>.42</td>
</tr>
<tr>
<td>Age at first E+P therapy (years)</td>
<td>53.4±0.37</td>
<td>52.8±0.27</td>
<td>.46</td>
</tr>
<tr>
<td>Duration of past E+P therapy (years)</td>
<td>5.7±0.27</td>
<td>5.5±0.58</td>
<td>.68</td>
</tr>
<tr>
<td><strong>Past oral contraceptive (OC) use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any past OC use (%)</td>
<td>569 (43.4)</td>
<td>122 (37.2)</td>
<td>.04</td>
</tr>
<tr>
<td>Age at first OC use (years)</td>
<td>29.0±0.30</td>
<td>28.8±0.57</td>
<td>.75</td>
</tr>
<tr>
<td>Duration of OC use (years)</td>
<td>4.9±0.21</td>
<td>5.3±0.47</td>
<td>.45</td>
</tr>
<tr>
<td><strong>Current HT use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any current HT&lt;sup&gt;bc&lt;/sup&gt; (%)</td>
<td>656 (53.9)</td>
<td>177 (55.5)</td>
<td>.61</td>
</tr>
<tr>
<td>E-alone therapy (%)</td>
<td>412 (34.2)</td>
<td>120 (38.0)</td>
<td>.21</td>
</tr>
<tr>
<td>E+P therapy (%)</td>
<td>210 (17.6)</td>
<td>51 (16.2)</td>
<td>.57</td>
</tr>
</tbody>
</table>

Notes: For categorical variables frequency and percentage are provided and for continuous variables, the mean and standard error are provided. P-values are obtained from a conditional logistic regression model to compare the two groups.

<sup>a</sup> Women could have received both “estrogen alone” and “estrogen plus progestin” therapies in the past. Therefore subcategories within the “any past HT use” entry are not mutually exclusive.

<sup>b</sup> Women could have received both “estrogen alone” and “estrogen plus progestin” therapies. Therefore the subcategories E-alone and E+P within the “any current HT use” entry are not mutually exclusive. Type of therapy was only completely available for HT trial participants, for the remainder of the women partial data on type of therapy was available. There were 104 women on whom current HT information was not available.
Table 5.3 Conditional Logistic Regression Multivariable Modeling for CLL/SLL

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past HT use (yes v. no)</td>
<td>1.317</td>
<td>1.014-1.710</td>
<td>.04</td>
</tr>
<tr>
<td>Past OC use (yes v. no)</td>
<td>0.729</td>
<td>0.557-0.955</td>
<td>.02</td>
</tr>
<tr>
<td>Obesity status (≥30 BMI vs. &lt;30 BMI)</td>
<td>0.729</td>
<td>0.549-0.968</td>
<td>.03</td>
</tr>
<tr>
<td>U.S. Region (^a)</td>
<td></td>
<td></td>
<td>.53</td>
</tr>
<tr>
<td>Northeast</td>
<td>1.280</td>
<td>0.906-1.808</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>1.183</td>
<td>0.839-1.668</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>1.081</td>
<td>0.755-1.549</td>
<td></td>
</tr>
<tr>
<td>Ever smoker (yes v. no)</td>
<td>0.995</td>
<td>0.780-1.270</td>
<td>.97</td>
</tr>
<tr>
<td>WHI Study Part (CT v. OS)</td>
<td>1.286</td>
<td>1.003-1.649</td>
<td>.05</td>
</tr>
<tr>
<td><strong>Model 2:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past E-alone use (yes v. no)</td>
<td>1.368</td>
<td>1.049-1.783</td>
<td>.02</td>
</tr>
<tr>
<td>Past E+P use (yes v. no)</td>
<td>1.259</td>
<td>0.942-1.681</td>
<td>.12</td>
</tr>
<tr>
<td>Past OC use (yes v. no)</td>
<td>0.733</td>
<td>0.559-0.959</td>
<td>.02</td>
</tr>
<tr>
<td>Obesity status (≥30 BMI vs. &lt;30 BMI)</td>
<td>0.728</td>
<td>0.548-0.968</td>
<td>.03</td>
</tr>
<tr>
<td>U.S. Region (^a)</td>
<td></td>
<td></td>
<td>.47</td>
</tr>
<tr>
<td>Northeast</td>
<td>1.303</td>
<td>0.921-1.844</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>1.201</td>
<td>0.851-1.695</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>1.088</td>
<td>0.759-1.560</td>
<td></td>
</tr>
<tr>
<td>Ever smoker (yes v. no)</td>
<td>0.993</td>
<td>0.778-1.267</td>
<td>.95</td>
</tr>
<tr>
<td>WHI Study Part (CT v. OS)</td>
<td>1.286</td>
<td>1.002-1.651</td>
<td>.05</td>
</tr>
<tr>
<td><strong>Model 3:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilateral oophorectomy (BOO) and past E-alone use (^b)</td>
<td></td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td>BOO and past E-alone</td>
<td>1.530</td>
<td>1.067-2.196</td>
<td>.02</td>
</tr>
<tr>
<td>BOO and no past E-alone</td>
<td>1.720</td>
<td>0.933-3.171</td>
<td>.08</td>
</tr>
<tr>
<td>No BOO and past E-alone</td>
<td>1.437</td>
<td>1.039-1.988</td>
<td>.03</td>
</tr>
<tr>
<td>Past E+P use (yes v. no)</td>
<td>1.287</td>
<td>0.961-1.724</td>
<td>.09</td>
</tr>
<tr>
<td>Past OC use (yes v. no)</td>
<td>0.757</td>
<td>0.575-0.997</td>
<td>.05</td>
</tr>
<tr>
<td>Obesity status (≥30 BMI vs. &lt;30 BMI)</td>
<td>0.692</td>
<td>0.516-0.928</td>
<td>.01</td>
</tr>
<tr>
<td>U.S. Region (^a)</td>
<td></td>
<td></td>
<td>.52</td>
</tr>
<tr>
<td>Northeast</td>
<td>1.290</td>
<td>0.910-1.830</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>1.114</td>
<td>0.784-1.583</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>1.057</td>
<td>0.734-1.521</td>
<td></td>
</tr>
<tr>
<td>Ever smoker (yes v. no)</td>
<td>0.988</td>
<td>0.771-1.265</td>
<td>.92</td>
</tr>
<tr>
<td>WHI Study Part (CT v. OS)</td>
<td>1.226</td>
<td>0.951-1.580</td>
<td>.12</td>
</tr>
</tbody>
</table>

Notes: BMI=Body mass index; BOO=Bilateral oophorectomy; CT=Clinical trials; E+P=Estrogen progestin therapy; E-alone=Estrogen therapy; HT=Hormone therapy; OC=Oral contraceptive; OS=Observational study. Conditional logistic regression was used to obtain the odds ratios and corresponding 95% confidence intervals for CLL/SLL. Interactions of all risk factors listed above with the past HT use variable (Model 1) and past E-alone use variable (Model 2) were evaluated. None of them were significant and therefore were not included in the final models. Similarly, there were no significant interaction terms in Model 3.  
\(^a\)West is reference group.  
\(^b\)Women with no BOO and no past E-alone is reference group.
Figure 5.1. Estrogen exposures and risk of CLL/SLL

CLL/SLL indicates chronic lymphocytic leukemia; BOO, bilateral oophorectomy; HT, hormone therapy; OC, oral contraceptive.
Chapter 6:  Pesticide Exposure and Incidence of Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma

6.1 Background

Although leukemia is a relatively rare form of cancer in adults (approximately 3% of all cancers), there are about 240,000 new cases diagnosed annually worldwide (43,800 in the U.S.) and approximately 200,000 deaths associated with it (23,300 in the U.S.) each year.\(^7\^8\) Leukemia is ranked fifth in person-years of life lost due to cancer, directly behind breast and pancreatic cancer.\(^9\) In addition, according to the National Cancer Institute’s Surveillance, Epidemiology and End Results (SEER) Program, during recent decades (between 1975 and 2010), there has been an increase among females in the U.S. of developing leukemia – with no clear explanation to support this trend.

In industrialized countries, Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma (CLL/SLL) is the most common type of adult leukemia and is a leukemia associated with aging- with causes largely unknown. Despite being a frequent type of leukemia, CLL/SLL is a relatively rare form of cancer (approximately 1% of all cancers). There are still about 15,720 new CLL/SLL cases diagnosed each year in the U.S – mainly in older adults (the lifetime risk of CLL/SLL is 0.52\%).\(^7\^8\^10\) In addition, according to the National Cancer Institute’s SEER Program, there has been an increased incidence of CLL/SLL between 1975 and 2010 among white females in the U.S.
Furthermore, though CLL/SLL is historically extremely rare among Asians, a rising incidence through a birth cohort effect has been noted in recent years. Asian populations born in the U.S. have the most notable increase in incidence, thereby supporting the notion that environmental risk factors are likely contributors to this disease.

CLL/SLL is almost exclusively a cancer of older people with median age of diagnosis at 70 years. Patients at this age have been subjected to several decades of DNA damage which necessitates the need to recognize the different means of exposure that lead to CLL/SLL. These exposures can be categorized into three different groups: 1) behavioral or lifestyle, 2) environmental, and 3) biological risk factors. Pesticide exposure spans both the behavioral and the environmental categories making it a very important risk factor to evaluate.

Pesticide, insecticide, and herbicide exposure are suspected risk factors for leukemia in general. These exposures originate from both agricultural applications (farming industry and private residences) as well as the manufacturing of these chemicals. Despite the heterogeneity in products, for a large proportion of chemicals used as pesticide and herbicides there is biological evidence of carcinogenicity. There have been several investigations published over the past couple of decades on this subject evaluating various chemicals either separately or in combination. However, these epidemiologic studies have not typically been powered adequately to detect significance in possible associations, and there is a lack of consistency in the characterization of exposure. Moreover, most of the research on this subject did not distinguish between
leukemia subtypes.

Certain pesticides, such as organophosphate pesticides and Alachlor, have been associated with increased leukemia incidence overall in a few studies. Thus, there seems to be initial epidemiologic evidence for pesticide exposure to be associated with leukemia incidence, however, no study to date has specifically evaluated pesticide exposure in relation to incidence of CLL/SLL in women. To investigate the association between past exposure and CLL/SLL we used data from the Women's Health Initiative (WHI), a large prospective study of post-menopausal women. We evaluated the association between adult lifetime pesticide exposure and the risk of CLL/SLL.

6.2 Methods

Study Design

The WHI was designed to address the major causes of morbidity and mortality in postmenopausal women and includes three clinical trials and an observational study. Details of the scientific rationale, eligibility requirements, and baseline participant characteristics of the WHI have been published elsewhere. Briefly, a total of 161,808 women, 50–79 years of age, were recruited at 40 clinical centers throughout the United States between September 1, 1993 and December 31, 1998. The WHI clinical trial includes four overlapping components: two hormone therapy trials (27,347 women), a dietary modification trial (48,835 women), and a calcium/vitamin D supplementation trial (36,282 women). Participants in the observational study (OS) included 93,676 women who were screened for the clinical trials but proved to be ineligible or unwilling
to participate or who were recruited through a direct invitation for the observational study. Women in the OS were administered additional questionnaires not completed by the clinical trial participants, which allowed for the collection of pesticide exposure data used in this study. The WHI study was overseen by institutional review boards at all 40 clinical centers and at the coordinating center, as well as by a study-wide data and safety monitoring board. All participants in the WHI gave informed signed consent and were followed up prospectively.

This study was a nested case control study (1:4 case-control matching using age and race) within the observational study arm of the WHI. The following participants were excluded from the original observational cohort of 93,676 for this analysis: 18,123 women who had a history of cancer at baseline, 12,506 women who had missing main exposure variables (either at baseline or at the year one follow-up form) and 359 women with a new leukemia diagnosis (other than CLL/SLL) during the study. This resulted in a sample size of 62,688 subjects, of whom 157 became confirmed CLL/SLL cases during the study. Applying a random selection of four controls for each case, matched by age and race, resulted in 628 controls and a total of 785 subjects for statistical analyses (after exclusions and matching).

**Measurements**

*Pesticide exposure measurement*

At the year one follow-up visit, participants were asked nine pesticide use-related questions with respect to past pesticide exposures: 1) Since age 21, have you or someone else ever poured, mixed, sprayed or applied insecticides (such as bug or flea spray,
garden/lawn/crop insecticides) in your immediate surroundings at home, leisure, or work? Sub-questions for participants marking “yes”: 2) Have you mixed insecticides? 3) Have you sprayed or applied insecticides? 4) Has lawn service applied insecticides? 5) Has commercial service applied insecticides? 6) Have you had other exposure to insecticides? 7) Specify years and times you mixed/applied insecticides. 8) Specify years and times lawn service applied insecticides. In addition, for pet owners only the following sub-question was asked: 9) Have you used any method to treat pets for fleas?

These questions were individually analyzed for their association with CLL/SLL. Furthermore, we derived the main exposure variable using the “location of exposure to insecticides” variable (i.e., “Since age 21, have you or someone else ever poured, mixed, sprayed or applied insecticides -such as bug or flea spray, garden/lawn/crop insecticides-in your immediate surroundings at home, leisure, or work?”). This newly derived variable called “any exposure to insecticides” was dichotomous and took on the value “yes” for exposures either at work, at home, or both versus the value “no” for no exposure at either location. To estimate cumulative exposure to pesticides, we derived a variable using the product of years mixed/applied and times mixed/applied variables collected in the questionnaire. The median value for each category for these two variables was used in the calculation to estimate cumulative exposure. Additionally, a potential surrogate variable for pesticide exposure was considered by analyzing the variable “Ever lived or worked on a farm” collected in the baseline observational study questionnaire.

Follow-up and ascertainment of cases

According to the World Health Organization (WHO), since 2008, SLL and CLL
have been considered one disease and one entity for disease classification, as malignant cells in both diagnoses exhibit the same immunophenotype. In our study we followed the most current guidelines (WHO 2008) which include SLL cases as CLL/SLL.

Incident CLL/SLL cases were identified by self-administered questionnaires (administered annually in the WHI clinical trial after 2005, and annually in the WHI observational study throughout the study), with all cases confirmed by medical record review. All CLL/SLL cases then were coded centrally in accordance with the Surveillance Epidemiology and End Results (SEER) coding guidelines. For these analyses, participants were followed up to CLL/SLL diagnosis, date of death, loss to follow-up, or end of WHI clinical trial or observational study follow-up, whichever occurred first.

**Statistical analysis**

Baseline characteristics were summarized for cases and controls displaying mean and standard error for continuous variables and frequencies and percentages for categorical variables. To compare the groups for variables not used in the matching, we used univariable conditional logistic regression models. Past pesticide exposure variables obtained at the one year follow-up were summarized similarly, as well as the surrogate pesticide exposure variable “Ever lived or worked on a farm?” Cumulative pesticide exposure was divided into quartiles and summarized accordingly. Univariable conditional logistic regression models were generated for CLL/SLL incidence for all collected pesticide use variables and other important risk factors. To evaluate important pesticide use related risk factors for CLL/SLL incidence in the presence of other
important risk factors, multivariable analysis was conducted. Interactions between pesticide use and other risk factors (i.e., obesity, smoking status) were evaluated by including a multiplicative interaction term in the regression model.

Conditional logistic regression models adjusted for important risk factors for the matched-pair cohorts, were used to estimate risk ratios of pesticide exposure for CLL/SLL incidence. In addition, all obtained risk ratio estimates, with their corresponding 95% confidence intervals and p-values, were calculated and reported. There were no adjustments made for multiple testing since only a few planned comparisons were made in this analysis. Analyses were performed using SAS 9.3 (SAS Institute, Cary NC).

6.4 Results

A total of 157 CLL/SLL cases meeting study inclusion criteria (i.e., no previous history of cancer, non-missing main exposure variables) were identified in the observational study, during a mean follow-up period of 12.7 years. Baseline characteristics according to case-control status are shown in Table 6.1. As age and race were both used in matching controls to cases, these variables were evenly distributed between the groups by definition. In regard to other important baseline characteristics, body mass index (BMI) tended to be lower for cases (P=.06) and when obesity status (i.e., BMI ≥ 30) was considered, cases were significantly less likely to be obese (P=.05). Smoking status (i.e., “Smoked at least 100 cigarettes ever”) was similar between cases and controls (P=.42). So was region of the U.S. they resided in (P=.95), and all other demographic factors. Data on pesticide exposure data since the age of 21 years according
to case-control status are presented in Table 6.2. When the location of pesticide exposure was considered (work, home, or both), cases tended to have higher exposures in all of these sub-categories (P=.06) compared to controls. Once these categories were collapsed into one “any exposure” category, cases had significantly higher rates of exposure, by about 10% (P=.01). In addition, seven measures of pesticide exposure were evaluated for subjects who reported any exposure since the age of 21, as was one measure for pet owners with respect to flea treatment. All of these sub-measures were similar between cases and controls. In addition, we found a significant association (P=.04) for increasing cumulative pesticide use to be associated with the cases compared to controls. Farm living/working history is a possible proxy for pesticide exposure, however, this was not confirmed in our study as the two measures seemed independent rather than closely correlated. Nearly the same percentage of women were exposed to pesticides with a history of farm living/working compared to those with no history of farm living/working (69% vs. 65%, P=.32) and farm living/working was similar between cases and controls.

To further describe the relationship between pesticide exposure and the risk for CLL/SLL, we estimated the odds ratio for CLL/SLL risk along with its 95% confidence interval. Based on univariable logistic regression, any pesticide exposure from the age of 21 resulted in an approximately 60% increase in the risk of developing CLL/SLL (odds ratio: 1.61, 95% confidence interval: 1.09-2.38, P=.01). This finding was confirmed by multivariable modeling (Table 6.3), which showed similar estimates for pesticide exposure as an important risk factor for CLL/SLL, after adjusting for other important risk factors and potential confounders, such as BMI, U.S. region, and smoking status.
In this nested case-control study, designed using a large prospective cohort of postmenopausal women, we found that women with past pesticide exposure have a significantly elevated risk of developing CLL/SLL. The association was observed for both work and home exposure and the type of pesticide application method (for example, lawn service, self-application, etc.) did not appear to have a further effect on the risk for CLL/SLL. The risk ratio of 1.65 we observed among these women is in the range of what has been found for CLL/SLL among men in a previous study (evaluating ever farmers vs. never farmers, OR=1.4) and in previous studies focusing on men. In addition, a recently published study on the same cohort of women from the WHI which evaluated all non-Hodgkin lymphomas in a cohort design setting, found a similar risk ratio for the CLL/SLL subset analysis.

We did not find an association with self-reported living or working on a farm and CLL/SLL incidence, as was found in other studies. The reason for this could be two-fold: 1) being on a farm either as one’s residence or place of employment, does not necessarily implicate contact with pesticides (only 69% of those on a farm reported pesticide exposure), and 2) farms are not the only locations where women were exposed to pesticides (65% of those not on a farm reported pesticide exposure). Therefore, this variable is not likely to be a precise measure of pesticide exposure relative to the main variable used in our study, which was self-reported pesticide exposure, undoubtedly a more objective measure. Most research to date that has attempted to evaluate specific pesticides used by
study participants (mainly farmers), has struggled with precisely isolating the specific types or categories of chemicals as people typically find it difficult to recall this level of information. Despite this challenge, certain pesticides such as the widely used class of organophosphates, have repeatedly been associated with leukemia in previous studies.\textsuperscript{215, 216, 217} It has been hypothesized that the biological mechanism by which organophosphates lead to leukemogenesis is by disturbing immune function, with permanent inhibition of acetylcholinesterase, an enzyme synthesizing acetylcholine into inactive metabolites, choline and acetate. Lymphocytes include key components of a cholinergic system, and prolonged acetylcholinesterase receptor stimulation, which could result from irreversible acetylcholine esterase inhibition, can alter lymphocytic activity.\textsuperscript{218} Although our study did not attempt to assess the particular type of pesticides the women were exposed to, we postulate that due to its common use, organophosphates were a large percentage of their exposure. Moreover, other frequently used pesticides and insecticides, such as pyrethrins (a class of pesticides derived from chrysanthemums and approved for use in organic farming) and nicotine have also been shown to be associated with leukemia, despite not yet having solid biological mechanisms explained. However, one might hypothesize that similar biological methods may apply which may be worthy of investigation.

Besides pesticide exposure being associated with an increased risk for CLL/SLL, our study further supports this finding by showing a dose-response type of relationship between quartiles of exposure (i.e., number of times exposed) and case-control status. This result strengthens the biological hypothesis that certain pesticides act as carcinogens and are leukemogenic.
Since pesticide use became prevalent starting in the 1950’s, the cohort in our study was the first generation of women at risk for exposure to these chemicals for a substantial period of time. These women were in their 20’s in the mid-twentieth century, thus not likely to have been exposed to pesticides during childhood, and therefore, we were limited to adulthood exposures only in our study. However, moving forward, newer studies should evaluate pesticide exposure starting in childhood as subsequent generations have potentially been affected earlier in life and children have an increased susceptibility to toxic chemicals. In addition, evaluating both synthetic and organic pesticides is important, as organic compounds can have high toxicity profiles as well.

In the U.S., the Environmental Protection Agency (U.S. EPA) and individual states regulate the use of pesticides, which involves registering and evaluating new pesticides, reviewing existing ones, and enforcing pesticide requirements. The U.S. EPA considers a range of research to evaluate cancer risk, including laboratory animal studies, metabolism studies, chemical relationship to other carcinogens, mode of carcinogenesis, and human epidemiological studies. Furthermore, since 1986 the U.S. EPA maintains a list of pesticides according to their hierarchical carcinogenetic potential with detail on type of cancer, ranging from “Carcinogenic to Humans” to “Evidence of Non-carcinogenicity for Humans”. The EPA’s Cancer Assessment Review Committee report is updated and released annually and provides an extensive list of studied pesticides and their cancer causing potential. During the past couple of decades since this list has been regularly published, it has already made an impact on minimizing non-essential pesticide applications (i.e., carcinogens or probable carcinogens) both in the agricultural industry and for lawn care. For example, as early as in 1972, after three
decades of extensive use, the first widely used pesticide (both in agriculture and around homes and gardens), DDT was banned by the U.S. EPA for a variety of reasons, its carcinogenic effect being one reason. The long-term epidemiological consequences of these regulations are yet to be evaluated. The hope would be that the incidence of associated cancers is lowered.

Our research has multiple strengths including 1) the nested case-control design that enabled us to efficiently assess the risk of CLL/SLL which is a rare cancer and is therefore difficult to study in a cohort setting, 2) its focus on women relative to pesticide exposure and association with CLL/SLL for the first time, 3) detailed data on exposures (including potential confounders) and on confirmed CLL/SLL diagnosis, and 4) use of a well-designed questionnaire to collect information on history of adult lifetime pesticide exposure by avoiding recall bias that would have likely resulted from asking chemical specific detail on pesticides. However, there are some possible limitations to our study such as pesticide use- our main exposure variable- was self-reported and it did not include detail on the specific chemicals used, nor did it assess blood/urine/tissue levels of pesticides, as well as some exposure misclassification was possible. In addition, there was no information collected on protective equipment use, which despite being a difficult variable to evaluate, could have been useful to adjust for in the models assessing risk for CLL/SLL.

In conclusion, postmenopausal women with pesticide exposures during their adult life had a significant increase in their risk for CLL/SLL. With the surge of pesticide use since the mid-twentieth century, this association could potentially explain the recent increase in CLL/SLL among women who have been at elevated exposure risk in both the
home and at their workplace. More studies are necessary to confirm these findings and to prospectively assess detailed information on pesticide exposure, as well as to explore additional biological mechanisms relative to these findings. If confirmed, pesticide exposure can be considered as a preventable risk factor for CLL/SLL and shed some light on possible causes for this hematological cancer of older people with essentially unknown risk profile thus far. As the average life expectancy of women continues to increase and is greater relative to men, aging related disease and their epidemiology, such as CLL/SLL were of great interest to further investigate in older women. Potential regulatory considerations and public health recommendations will also need to be addressed.
Table 6.1 Baseline Characteristics by Control and Case Status

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=628)</th>
<th>Cases (n=157)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>63.3±0.3</td>
<td>63.3±0.5</td>
<td>Matched</td>
</tr>
<tr>
<td>Race</td>
<td>Matched</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American (%)</td>
<td>44 (7.0)</td>
<td>11 (7.0)</td>
<td></td>
</tr>
<tr>
<td>White (%)</td>
<td>564 (89.8)</td>
<td>141 (89.8)</td>
<td></td>
</tr>
<tr>
<td>Other (%)</td>
<td>20 (3.2)</td>
<td>5 (3.2)</td>
<td></td>
</tr>
<tr>
<td>U.S. Region</td>
<td></td>
<td></td>
<td>.95</td>
</tr>
<tr>
<td>Northeast (%)</td>
<td>149 (23.7)</td>
<td>34 (21.7)</td>
<td></td>
</tr>
<tr>
<td>South (%)</td>
<td>155 (24.7)</td>
<td>40 (25.5)</td>
<td></td>
</tr>
<tr>
<td>Midwest (%)</td>
<td>149 (23.7)</td>
<td>37 (23.6)</td>
<td></td>
</tr>
<tr>
<td>West (%)</td>
<td>175 (27.9)</td>
<td>46 (29.3)</td>
<td></td>
</tr>
<tr>
<td>Body mass index (BMI), kg/m²</td>
<td>27.3±0.2</td>
<td>26.5±0.4</td>
<td>.06</td>
</tr>
<tr>
<td>Obese (% BMI ≥30)</td>
<td>163 (26.0)</td>
<td>29 (18.5)</td>
<td>.05</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>Never smokers (%)</td>
<td>338 (54.2)</td>
<td>79 (50.6)</td>
<td></td>
</tr>
<tr>
<td>Former smokers (%)</td>
<td>239 (38.4)</td>
<td>69 (44.2)</td>
<td></td>
</tr>
<tr>
<td>Current smokers (%)</td>
<td>46 (7.4)</td>
<td>8 (5.1)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 6.1 Baseline Characteristics by Control and Case Status (continued)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=628)</th>
<th>Cases (n=157)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education, college graduate or above (%)</td>
<td>278 (44.6)</td>
<td>73 (46.5)</td>
<td>.66</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td>.31</td>
</tr>
<tr>
<td>Never married (%)</td>
<td>27 (4.3)</td>
<td>2 (1.3)</td>
<td></td>
</tr>
<tr>
<td>Divorced or separated (%)</td>
<td>87 (13.9)</td>
<td>21 (13.4)</td>
<td></td>
</tr>
<tr>
<td>Widowed (%)</td>
<td>100 (16.0)</td>
<td>27 (17.2)</td>
<td></td>
</tr>
<tr>
<td>Presently married (%)</td>
<td>398 (63.8)</td>
<td>106 (67.5)</td>
<td></td>
</tr>
<tr>
<td>Marriage-like relationship (%)</td>
<td>12 (1.9)</td>
<td>1 (0.6)</td>
<td></td>
</tr>
<tr>
<td>Family income</td>
<td></td>
<td></td>
<td>.96</td>
</tr>
<tr>
<td>&lt;$10,000-$19,999 (%)</td>
<td>80 (13.5)</td>
<td>19 (12.7)</td>
<td></td>
</tr>
<tr>
<td>$20,000-$49,999 (%)</td>
<td>245 (40.6)</td>
<td>58 (38.7)</td>
<td></td>
</tr>
<tr>
<td>$50,000-$99,999 (%)</td>
<td>195 (32.3)</td>
<td>51 (34.0)</td>
<td></td>
</tr>
<tr>
<td>≥$100,000 (%)</td>
<td>84 (13.9)</td>
<td>22 (14.7)</td>
<td></td>
</tr>
<tr>
<td>Health insurance (yes) (%)</td>
<td>596 (95.8)</td>
<td>152 (97.4)</td>
<td>.35</td>
</tr>
</tbody>
</table>
### Table 6.2. Pesticide Exposure by Control and Case Status

<table>
<thead>
<tr>
<th>Variables</th>
<th>Controls (n=628)</th>
<th>Cases (n=157)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any pesticide exposure</td>
<td>403 (64.2)</td>
<td>117 (74.5)</td>
<td>.01 (.04 b)</td>
</tr>
<tr>
<td>Location of exposure</td>
<td></td>
<td></td>
<td>.06</td>
</tr>
<tr>
<td>Work only (%)</td>
<td>14 (2.2)</td>
<td>6 (3.8)</td>
<td></td>
</tr>
<tr>
<td>Home/leisure only (%)</td>
<td>323 (51.4)</td>
<td>89 (56.7)</td>
<td></td>
</tr>
<tr>
<td>Work and home/leisure (%)</td>
<td>66 (10.5)</td>
<td>22 (14.0)</td>
<td></td>
</tr>
<tr>
<td>No exposure (%)</td>
<td>225 (35.8)</td>
<td>40 (25.5)</td>
<td></td>
</tr>
<tr>
<td>Mixed (%)</td>
<td>91 (22.8)</td>
<td>25 (21.4)</td>
<td>.74</td>
</tr>
<tr>
<td>Sprayed or applied (%)</td>
<td>268 (67.2)</td>
<td>71 (60.7)</td>
<td>.19</td>
</tr>
<tr>
<td>Lawn service (%)</td>
<td>126 (31.6)</td>
<td>35 (29.9)</td>
<td>.73</td>
</tr>
<tr>
<td>Commercial service (%)</td>
<td>134 (33.6)</td>
<td>40 (34.2)</td>
<td>.90</td>
</tr>
<tr>
<td>Other exposure (%)</td>
<td>47 (11.8)</td>
<td>18 (15.4)</td>
<td>.30</td>
</tr>
<tr>
<td>Years mixed/applied</td>
<td>2.7 ±0.1</td>
<td>2.6 ±0.2</td>
<td>.43</td>
</tr>
<tr>
<td>Times mixed/applied</td>
<td>0.8 1±0.04</td>
<td>0.72 ±0.07</td>
<td>.36</td>
</tr>
<tr>
<td>Pets treated for fleas (%)</td>
<td>76 (15.3)</td>
<td>19 (15.3)</td>
<td>.99</td>
</tr>
<tr>
<td>Ever lived/worked on a farm (yes) (%)</td>
<td>169 (26.9)</td>
<td>44 (28.0)</td>
<td>.78 (.32 e)</td>
</tr>
</tbody>
</table>

a Pesticide exposure the since age 21, irrespective of location.
b Pesticide exposure the since age 21.
c Sub-question for subjects with any pesticide exposure the since age 21.
d Sub-question for subjects who lived with pets since age 21.
e 69% of those living/working on a farm were exposed to pesticides v. 65% of those not living/working on a farm (P=.32), indicating the measures are independent.
Table 6.3 Conditional Logistic Regression Multivariable Model for CLL/SLL

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any exposure (yes v. no)</td>
<td>1.65</td>
<td>1.11-2.45</td>
<td>.01</td>
</tr>
<tr>
<td>Obesity status (≥30 BMI vs. &lt;30 BMI)</td>
<td>0.65</td>
<td>0.42-1.01</td>
<td>.06</td>
</tr>
<tr>
<td>U.S. Region (West is reference group)</td>
<td></td>
<td></td>
<td>.99</td>
</tr>
<tr>
<td>Northeast</td>
<td>0.92</td>
<td>0.55-1.55</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>1.01</td>
<td>0.61-1.68</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>1.00</td>
<td>0.61-1.66</td>
<td></td>
</tr>
<tr>
<td>Ever smoker</td>
<td>1.18</td>
<td>0.83-1.68</td>
<td>.37</td>
</tr>
</tbody>
</table>

Note: Conditional logistic regression was used to obtain the odds ratios and corresponding 95% confidence intervals for CLL/SLL. Interactions of all risk factors listed above with the “any exposure” variable were evaluated. None of them were significant and therefore were not included in the final model.
Chapter 7. Summary, Conclusions, Implications

7.1 Summary

The overall goal for this study was to discover novel risk factors for Chronic Lymphocytic Leukemia and Small Lymphocytic Lymphoma (CLL/SLL) that are potentially modifiable with behavior or lifestyle changes. This primary objective was motivated by the fact that, presently, the only well-established risk factors for this most frequent adult leukemia in the western world are all inherited (i.e., not modifiable). This study focused on specific personal habits and also on pesticide exposures, in order to see if there are any associations that are unique for industrialized countries. In addition, our study had the secondary objective to try to understand why the rate of CLL/SLL in women is have half the rate in men. This second objective prompted us to focus on women in order to reveal risk factors unique to them, with particular attention to endogenous and exogenous estrogens and hence we explored various means of estrogenic exposures relative to CLL/SLL risk that were possible since we used data from the WHI.

Employing an age and race matched nested case-control study
design using the WHI, allowed us to investigate CLL/SLL risk and its relationship, in postmenopausal women, to the main objectives of this research. Briefly, we evaluated: in Aim 1) personal habits; in Aim 2) hormonal exposures; and in Aim 3) pesticide exposures. Besides permitting us to focus on the female population in detail and providing an extensive data base of potential risk factors, the WHI enrolled women who were of an ideal age to investigate our objectives and specific aims since they were post-menopausal and the average age of CLL/SLL onset is at around 70 years. In addition, the large sample size of 161,808 women in the WHI, resulted in a relatively large number of confirmed CLL/SLL cases compared to other similar studies. This allowed us comprehensively evaluate the risk factors under consideration.

Our finding relative to pesticide exposures provides plausibility for our primary objective of why CLL/SLL is significantly more prevalent in industrial countries compared to developing countries, as we found exposures to pesticides to be an adverse risk factor and it is recognized that pesticides use is three-fold higher in industrialized countries relative to the rest of the world. With respect to our secondary objective of investigating female-specific exposures, we discovered various estrogenic associations that uncover possible mechanisms by which women have lower rates of CLL/SLL compared to men. The risk factors we found are thought to
increase estrogen in the body either through endogenous mechanisms (such as the result of obesity) or through exogenous sources (such as from OC use or coffee consumption). In addition, we postulated that a lack of adequate estrogen in women who underwent BOO is an adverse risk factor for CLL/SLL. This was indicated by our finding that women who took E-alone therapy with the common indication of BOO, were at an increased risk for CLL/SLL despite their therapy, most likely due to still not meeting their optimal estrogen needs.

7.2 Conclusions

Aim 1

Our first important result was that, with respect to drinking habits, we found coffee, a widely used and popular beverage, to be protective against CLL/SLL, when consumed regularly in the CT participants of the WHI. Relative to alcohol consumption, women in our study reported very low alcohol drinking (less than 2.5 drinks per week on average) and therefore our analyses may not have had enough variability in the data to be able to detect any potential differences between cases and controls. Finally, we were not able to detect any association between diet or exercise on the risk of CLL/SLL. These null findings are generally consistent with the inconclusive and mainly negative results of smaller studies by other groups. With respect to exercise, the lack of a significant finding is in line with the
fact that obesity offers protection for CLL/SLL, and as obese women had lower MET hours per week, by about an hour and a half compared to non-obese women, the evaluation of exercise is confounded by body weight.

Aim 2

We found that behaviors leading to higher estrogen levels are protective against CLL/SLL. These findings were consistent with certain biological hypotheses that show estrogens be protective against cancers. Specifically, in our second aim we found that women who were obese at baseline and/or were users of OCs in the past had a significant reduction in their CLL/SLL risk. On the other hand, past E-alone therapy did not offer protection; on the contrary, it was seen to be associated with increased CLL/SLL risk. We postulated this finding to be a result of low levels of estrogen- despite therapy- among women who underwent BOO. Overall, since we found high estrogen levels protective and low levels adverse, our investigation of hormone exposures (endogenous and exogenous) yielded in a consistent discovery and an important epidemiological finding that will need to be investigated further for biological mechanisms in laboratory and clinical settings. Moreover, this finding revealed a very plausible mechanism as to why women have a much lower (i.e., half) rate of CLL/SLL compared to men who do not have the protection from estrogen.
Aim 3

When investigating our third aim, we showed that pesticide exposures during a woman’s adult life significantly increase the risk of CLL/SLL. This adverse impact was independent of other potentially important risk factors, such as region of the U.S. and smoking (which was not a predictor of CLL/SLL risk), and was true for both in the home and work pesticide exposures and is consistent with a recent study conducted using WHI data evaluating all NHLs. Although certain pesticides are known to be estrogenic, we postulated that their primary mechanism by which they increase the risk of CLL/SLL is through damaging DNA and thereby causing chromosomal aberration that in turn lead to CLL/SLL. We concluded that since pesticide use is significantly more prevalent in industrialized countries (relative to developing countries) and they bear an adverse impact, this could, at least in part, explain why CLL/SLL is mainly a hematological malignancy of the western world.

In order to comprehensively evaluate the three aims, we sought to simultaneously assess the main risk factor variables from each individual aim by multivariable modeling. Since pesticide exposure was only available for women on the OS, cell sizes for this analysis were inadequate and this analysis was not feasible.
7.3 Limitations, Strengths, and Implications

Overall, the chief limitations to our study pertain to the nature of exposure data collection, which was collected retrospectively through self-reported questionnaires. For example, although the WHI employed a well-designed and validated food frequency questionnaire, these data tend to be difficult to recall and hence could lead to recall bias. In addition, due to pesticide data only being available for about half of the study participants (i.e., OS arm), we did not have adequate power to evaluate all three of our aims together in one model. However, future, well-powered studies might be able to examine these risk factors given our findings on each one. Furthermore, data validity with respect to alcohol use may have been compromised, as this behavior is not generally regarded as socially desirable. With respect to our third aim, detail on the type of pesticide chemicals used was not captured which only allowed for more general conclusions.

This study had multiple strengths. Most importantly, our results have some important public health implications for understanding risks for CLL/SLL, which were not clear before our study. Several of our findings call for further research into better understanding the underlying biological mechanisms. Existing literature provided some support for a relationship
for exogenous estrogens and NHL risk overall, but CLL/SLL has had very limited study. With the large number of confirmed CLL/SLL cases in the WHI we were able to conduct an adequately powered epidemiologic investigation of postmenopausal women and distill sex-specific risk factors that could elucidate why women have lower rates of CLL/SLL than men. Consequently, we showed that behaviors leading to higher endogenous and exogenous estrogen levels protect women against the most common type of adult leukemia, CLL/SLL. Thus, we proposed that estrogen plays an integral role in reducing the incidence of CLL/SLL in women compared to men. However, it is important to emphasize that despite identifying obesity as a protective risk factor for CLL/SLL, this finding should only be used to better understand CLL/SLL etiology as the morbidity associated with obesity outweighs any of these benefits.

The other important public health implication is that coffee consumption appears to offer protection against CLL/SLL. Coffee is known to contain a mixture of caffeine and polyphenols and with the rising worldwide popularity of coffee drinking, research investigating the bioactivity of these various compounds has been of increasing interest. Our study was the first to focus on evaluating coffee drinking specific to CLL/SLL risk, and we discovered a weak relationship between higher coffee drinking and lower risk for CLL/SLL.
The final public health implication of our study is that pesticide exposure increases the risk of CLL/SLL. This finding is well in concordance with published data by the EPA on the carcinogenic potential of numerous pesticides (detailed according weight of evidence), some of which have been discontinued in recent decades - but still were exposures during the lifetime of the WHI women - and some of which are still being used widely. In order to provide specific guidance as to which particular pesticides should be avoided for CLL/SLL prevention, additional studies that measure chemical exposures would need to be done. Minimizing pesticide exposure in general, is likely to be beneficial for risk reduction.

The findings of our study provide new information on CLL/SLL risk factors that are potentially modifiable and also help explain why CLL/SLL incidence is higher in industrialized countries as well as in men. Until now, there was very little known relative to epidemiological associations. We hypothesized that hormone exposures would have an effect on risk but the direction of the association was not clear upfront and, in fact, we expected an adverse impact of estrogen for our outcome of interest. Upon completion of our study, the protective effect of estrogen was revealed which consequently allowed us to directly use this finding to elucidate why women in general have a lower incidence of CLL/SLL compared to men. Our results also provide evidence for coffee drinking to
be a protective, which was also an unanticipated finding. In addition, relative to dietary habits, alcohol use, and physical activity, we were unable to pick up any associations, which was unforeseen, nevertheless consistent with other research. Concerning pesticides, our findings were in line with our hypotheses and with the other recently done study using the WHI cohort by Schinasi et al., and we showed a significant increase in CLL/SLL risk due to this environmental exposure. As pesticides are widely applied in industrialized countries, their exposure and the associated CLL/SLL risk can potentially affect a large percentage of women and could also explain the recent trend in the increase of CLL/SLL incidence among women. Future epidemiologic studies could aim to further refine our results by prospectively evaluating exposures with particular focus on estrogenic factors. In addition, laboratory and clinical studies are needed to confirm the mechanisms of important exposures we discovered, both in vitro and in vivo.
Appendix: Women’s Health Initiative Data Distribution Agreement
WHI DATA DISTRIBUTION AGREEMENT
MEMORANDUM

Date: November 26, 2013
To: Investigators seeking access to WHI data
From: WHI Clinical Coordinating Center
Subject: Accessing WHI data sets

To obtain access to the WHI data available on the WHI Study Operations website, please send the attached forms to the Help Desk by fax (206-667-4142) or email (helpdesk@whi.org).

When completing the Data Distribution Agreement, please be aware of the following:

- You must sign as the "Recipient".
- The WHI Principal Investigator who is sponsoring your manuscript or ancillary study (AS) must sign the attached form as the "sponsoring PI".
- You need to indicate the number of the approved manuscript proposal, BAA, or AS for which you intend to use the WHI data.
- If you will be using WHI data for a BAA or AS, please confirm you have submitted the study's data to the CCC; you are required to do so before obtaining access to WHI data.
- If you will be working with an analyst or other collaborator who will be using the data, that collaborator and you will need to complete the Data Distribution Agreement for Use of Data by Collaborators, which is the 2nd page of the attached form.

Once the signed agreement is received and we verify you have met the requirements for data access, we will send you an email containing a username and password that can be used to access the data for 3 months.

You must download the data within this 90-day period; extensions will not be granted. Because of this strict time limit, please download all the data files you might possibly need for your analyses.

Please note that this material is meant to be self-explanatory. The data are provided in fixed-width, space-delimited ASCII format. The data documentation and basic SAS code, which may be used for reading the data, are provided in the downloaded files. CCC documentation for your use is limited to assuring access from appropriately configured machines and providing additional documentation needed to use these data. Please send such queries to the email address: helpdesk@whi.org.

Best wishes for productive use of these data,

WHI Clinical Coordinating Center
WHI DATA DISTRIBUTION AGREEMENT

TERMS: “Recipient” is the investigator who is requesting a copy of the data set. The “data set” refers to any portion, part or subset of the data files available on WHI Study Operations website (www.whiops.org). “Sponsoring PI” is the WHI Principal Investigator overseeing the Recipient in his/her WHI work.

RECIPIENT INFORMATION

NAME Kati Maharry
INSTITUTION The Ohio State University

EMAIL maharry.2@osu.edu
POSITION / TITLE Lead Research Specialist

NAME OF SPONSORING P.I.
(Not required if the Recipient is a BAA PI) Rebecca Jackson, MD

APPROVED MANUSCRIPT #(/s) Ms2273v2
APPROVED AS / BAA #

AGREED TERMS AND CONDITIONS

I, the Recipient, have reviewed carefully and fully understand the terms for use of the WHI distributed data set, which are briefly listed below. I agree to abide fully by these terms and accept full responsibility for use and protection of the data set at my institution.

1. The data set may only be used for approved analyses in the manuscript number(s) listed above.
2. The data set may not be shared with outside commercial enterprises.
3. Under my supervision, investigators at my institution may use the data set for the purposes specified above and in compliance with all WHI policies (available on www.whi.org). All investigators who use the data set will review the terms of use and sign the Collaborator Data Distribution Agreement attached hereto as Attachment A.
4. Qualified individuals who sign Attachment A will provide statistical consultation for the analyses. Information on analytic methods used will be provided to the P&P Committee (according to the policies above) for review prior to publication of manuscripts. A biosketch may be used to document competence in conducting analyses.
6. I will not use the data set either alone or in conjunction with any other information in any effort whatsoever to identify participants.

SIGNATURE OF RECIPIENT Kati Maharry DATE 3/25/14

SIGNATURE OF SPONSORING PI (Not required if the Recipient is a BAA PI)

☐ Check if Requesting Temporary Access to the WHI Sharepoint Website

WHI CLINICAL COORDINATING CENTER
1100 Fairview Ave. N., PO Box 19024, M3-A410, Seattle, Washington 98105-1024 Fax: 206-667-4142

162
**Bibliography**


http://www.lls.org/content/nationalcontent/resourcecenter/freeeducationmaterials/generalcancer/pdf/facts.pdf

http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/leukaemia-key-facts/uk-leukaemia-statistics

10 http://www.lls.org/#/diseaseinformation/leukemia/chroniclymphocyticleukemia/incidence/


16 Strom SS, Yamamura Y, Kantarijian HM, Cortes-Franco JE. Obesity, Weight Gain, and Risk of Chronic Myeloid Leukemia. Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2010 August 9. Published in final edited form as: Cancer Epidemiol Biomarkers Prev. 2009 May; 18(5): 1501–1506. doi: 10.1158/1055-9965.EPI-09-0028. PMCID: PMC2918285


46 National Cancer Institute: Obesity and Cancer Risk. A fact sheet that summarizes research on the potential link between obesity and cancer risk. 2012.


57 http://www.trecscience.org/trec/


Mahmud SM, Franco EL, Aprikian AG. Use of nonsteroidal anti-inflammatory


88 Ross JA, Sinner PJ, Blair CK, Cerhan JR, Folsom AR. Hormone therapy is not associated with an increased risk of leukemia (United States). Cancer Causes


97 Smedby KE, Hjalgrim H, Melbye M, Torrång A, Rostgaard K, Munksgaard L,


1984836.


139 The Severe Chronic Neutropenia International Registry: https://depts.washington.edu/registry/


158 https://www.whi.org/data/Pages/categories/demographics.aspx

159 http://www.accessdata.fda.gov/scripts/cder/ndc/dsp_searchresult.cfm


213 Brown LM; Blair A; Gibson R; Everett GD; Cantor KP; Schuman LM; Burmeister LF; Van Lier SF; Dick F. Pesticide exposures and other agricultural risk factors for leukemia among men in Iowa and Minnesota. Cancer research, 1990 Oct 15; 50(20): 6585-91


219 http://www2.epa.gov/osa/guidelines-carcinogen-risk-assessment

220 http://www.epa.gov/pesticides/foia/

221 http://www2.epa.gov/aboutepa/ddt-ban-takes-effect