2016

The Association of Cancer Development in Patients with Systemic Lupus Erythematosus

Rose Michelle Coley
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral dissertation by

Rose Coley

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Aaron Mendelsohn, Committee Chairperson, Public Health Faculty
Dr. Diana Naser, Committee Member, Public Health Faculty
Dr. Roland Thorpe, University Reviewer, Public Health Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2016

Abstract

The Association of Cancer Development in Patients with Systemic Lupus Erythematosus

by

Rose Michelle Coley

MPH, Walden University, 2011

BS, University of Mount Olive, 2008

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

March 2016

Abstract

Both cancer and autoimmune diseases have been associated with numerous factors that may independently lead to the development of either disease. When these factors overlap the difficulty in assessing association is compounded. The numerous factors that are thought to cause systemic lupus erythematosus (SLE), which leads to the development of cancer, makes the study of an association between the 2 diseases challenging. The purpose of this study was to examine whether the risk of cancer development increased in SLE patients compared to the risk in non-SLE patients. Researchers have not shown consistent relationships of cancer development in patients with SLE; however, consideration of the various factors that contribute to the diseases is necessary to measure an association between the 2 diseases. This study used the Clinical Practice Research database (CPRD), a large, population-based database to test the relationship between SLE and cancer. A matched retrospective cohort study among SLE (n=3025) and non-SLE (n=180555) patients was conducted using the propensity score methodology to help balance the differences between the comparison groups. The propensity score methodology created a similar distribution of observed baseline covariates between the 2 groups. With adjustment for age, the predictor variable of SLE indicates that a patient with SLE is still 2.7 times more likely to develop cancer than is a non-SLE patient. The study outcomes could promote positive social change by reinforcing current recommendations for cancer screenings in persons with SLE, which could enhance the ability to detect cancer early enough to decrease mortality because of cancer in persons with SLE.

The Association of Cancer Development in Patients with Systemic Lupus Erythematosus

by

Rose Michelle Coley

MPH, Walden University, 2011

BS, University of Mount Olive, 2008

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

March 2016

Dedication

I dedicate this dissertation to my Mother Eula Taylor Coley and my Grandmother Sarah Green Coley. Your examples taught me to represent myself well at all times. You both made me believe that I can achieve whatever I want to as long as I work for it and never give up. If only you could be here to witness this.

I also dedicate this dissertation to my children and grandchildren, Taylor, Uriah, Ezekiel, and Nasiir. You all have witnessed me going through this dissertation process, and I did not give up because I hope that you all will also realize that each of you can accomplish your goals if you set them and work diligently towards meeting them. My expectation is that each of you will attain higher heights than your parents/grandparents.

Acknowledgments

I thank the faculty of Walden University, family members, and friends who have helped me reach this point in my academic career. I also thank my colleagues who served as mentors throughout this dissertation process.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Introduction

Systemic lupus erythematosus (SLE) has accounted for a significant number of deaths because of the devastation that this disease causes on numerous organ systems of the body. SLE disease activity is better understood and is controlled with the use of medications (Liang et al., 2012). The control of SLE flares has resulted in decreased deaths due to SLE activity, and people affected by the disease are living longer that they did 25 years ago (Liang et al., 2012). However, the increase in the lifespan of people with SLE has led researchers to determine that other chronic diseases are often the cause of death for persons with SLE (Liang et al., 2012).

People with SLE are more likely to be diagnosed with certain types of cancer as compared to persons without SLE (Bernatsky et al., 2005; Kiss, Kovacs, & Szodoray, 2010; Parikh-Patel, White, Allen, & Cress, 2008). Several risk factors have been proposed to account for the increased development of cancer among persons with SLE, including the use of immunosuppressive agents that increases the development of malignancies (Bernatsky, Ramsey-Goldman, & Clarke, 2006). Other risk factors that may increase the incidence of cancer in people with SLE include genetic predisposition, lifestyle-related risk factors, and abnormalities in cell death regulation (Kiss et al., 2010; Tincani, Taraborelli, & Cattaneo, 2010; Schultz & Harrington, 2003). In this study, I assessed the rate of cancer in patients with SLE compared to patients without SLE.

The results of this study may provide a better understanding of cancer development in people with SLE. The results did reveal a greater association of cancer development in patients with SLE as compared to non-SLE patients. The study outcomes could promote positive social change by reinforcing current recommendations for cancer screenings particularly in persons with SLE. Stronger adherence to cancer screening recommendations could enhance the ability to detect a cancer early enough so that treatment can be implemented that may result in a higher likelihood of eradicating the cancer and decreasing mortality due to cancer in persons with SLE. The results from this study may also equip persons with SLE with scientifically based knowledge that may enable them to make decisions regarding their care with a clearer understanding of the cancer risks inherent to persons with SLE, particularly when factored with their knowledge of their personal familial risks for cancer development.

In this chapter, literature related to cancer development in persons with SLE is summarized to identify gaps in current knowledge. The significance of whether an association in cancer development exists in persons with SLE is presented. A discussion of the purpose of the study, research questions, and the theoretical framework for the study are presented. Both dependent and independent variables and their definitions are provided. The research problem addressed and the limitations of the study are described in this section. The chapter concludes with potential contributions that this study will make toward the body of scientific knowledge, medical practice, and positive social change.

Background

Several risk factors for the development of cancer among persons with SLE have been proposed that would support an increased risk of cancer in this population, but these risks have not been fully explored. Both SLE and cancer have etiologic agents in common. In assessing the association of cancer in autoimmune diseases, many of the treatments used for the treatment of autoimmune diseases have been associated with the development of cancer (Azab et al., 2008). In addition, both cancer and autoimmune diseases have been associated with diet, air quality, exposure to certain drugs, and personal habits, which makes the study of an association between the two diseases challenging. Bei, Masuelli, Palumbo, Modesti, and Modesti (2009) concluded that the interaction between autoantibodies in cancer patients and autoimmune patients have similar antibodies that must be considered because they can change the properties of each other and impact the growth and progression of each disease. SLE is an autoimmune disease that results when the body starts to produce antibodies against itself. Signaling between antibodies (T cells and B cells) is impaired. Both T cells and B cells are instrumental in the development of several cancers. The interaction of these two antibodies must be considered because the properties of each are interdependent, as concluded by Bei et al. (2009).

Researchers have not consistently shown positive relationships between cancer development and SLE. Some scholars have shown that an increased risk of hematological cancers in SLE patients is likely because of other causes such as medications (Bernatsky et al., 2008). An understanding of the effects of the various factors that contribute to a

disease is necessary to measure an association of the variable to disease development (Broadbent, 2009). Because autoimmune diseases, such as SLE, have several risk factors that are thought to cause the disease, a more definitive association of the proposed causes of SLE are needed to measure better the association of the SLE risk factors in cancer development. The use of the high dimensional propensity model in this study balanced the two cohorts, so that a better assessment of the selected risk factors can be conducted.

Previous studies conducted have concluded that cancer rates are increased in SLE patients. The previous studies have shown variability in which types of cancers have been found to be increased. Kiss et al. (2010) identified several malignancies, such as non-Hodgkin's lymphoma (NHL), cervical cancer, and bronchial carcinomas, that were found to be increased in SLE patients, with the highest risk occurring in the first year of disease diagnosis. Nived, Bengtsson, Jönsen, Sturfel, and Olsson (2001) followed SLE patients to determine the rate of new malignancies following initial SLE diagnosis. Parikh-Patel et al. (2008) proposed that patients with SLE have an increased risk of developing hematologic, kidney, and thyroid cancers. Hildalgo-Conde et al. (2013) suggested that the incidence of cancer was four times greater than expected in a cohort study of Spanish patients. As shown by the findings from these studies, there is variability in the types of cancers that have been found to have an increased incidence in SLE patients.

Most studies conducted to assess the association of cancer development in patients with SLE have been relatively small in size; therefore, the researchers have been unable adequately to examine the development and exposures of cancers, especially those that are not common (Bernatsky et al., 2008). In this study, consideration was given to

the multiple factors that play a role in the development of SLE and in cancer. A large, population-based database was used to assess the relationship between SLE and cancer in a large population. The larger sample size in this study allowed cancers that occur at a lower incidence to be examined. Finally, I examined the incidence rates of various cancer types found in patients with SLE. The design of the study and the methods to assess the findings were dynamic enough to allow for a more detailed analysis that could be easily understood.

Researchers have shown inconsistent relationships between cancer development and SLE. The positive associations between hematologic cancers and several autoimmune diseases have been demonstrated in several studies. Bernatsky et al. (2008) studied immunosuppressive therapy in SLE patients and found that immunosuppressive therapy may not be the principle driving factor for overall cancer risk, but may contribute to an increased risk of hematological cancers and is likely a plethora of causes that result in cancer development in persons with SLE. Kiss et al. (2010) identified that several malignancies, such as NHL, cervical cancer, and bronchial carcinomas, are found at a higher rate in patients with SLE; in addition, SLE patients have an increased incidence and risk of cancer development, with the highest risk occurring in the first year of disease diagnosis. Nived et al. (2001) followed SLE patients to determine the rate of new malignancies following initial SLE diagnosis. Interestingly, Nived et al. did not find an overall increase in the cancer in SLE patients but did find increased frequencies of NHL.

Although the etiology of autoimmune disorders is not known with certainty, several hypotheses have been proposed to explain the reason for the body to begin

attacking itself. Nakazawa (2008) described environmental triggers, such as chemicals, drugs, hormones, infections, stress, hormones, behaviors, and diet that are thought to cause the body to start producing antibodies against itself. Because a high number of autoimmune diseases, such as SLE, are more prevalent in females, hormones may contribute to the development of autoimmunity. Approximately 78% to 90% of people with SLE are women (Fairweather & Rose, 2004). Because such a high percentage of women who develop SLE are women, the activities of hormones, such as estrogen, a hormone that has higher levels in women, should be acknowledged in an assessment of cancer development in SLE patients. Estrogen is also considered a risk factor for cancers such as breast cancer.

Immunosuppressive therapy is a common treatment for SLE. This therapy is implemented to suppress the immunology system and decrease a reaction to the production of the antibodies that the body is producing against itself. Bernatsky et al. (2008) proposed that immunosuppressive medications used to treat SLE may not be the main factor for overall cancer risk in SLE patients. Immunosuppressive medications may contribute to an increased risk of developing hematologic cancers, and future scholars should evaluate other factors that increase the risk for malignancy in persons with SLE. Chang et al. (2005) found that the use of nonsteroidal anti-inflammatory drugs (NSAIDs) to treat inflammation increased the risk of NHL. Engels et al. (2005) found Sjögren's syndrome to have the strongest association with immune-related conditions and immune modulating medications as risk factors for NHL. Although SLE is different from Sjögren's syndrome, they are both autoimmune diseases and have some similarities.

Findings from studies of other autoimmune diseases, such as Sjögren's syndrome, must be considered when studying SLE as in this study because of the similar risk factors and treatments.

There are variable relationships between cancer developments in patients with SLE (Bernatsky et al., 2008; Kiss et al., 2010; Parikh-Patel et al., 2008). The smaller sample sizes of previous studies have not allowed for an examination of cancers that are diagnosed at a lower incidence, which is important because of the various associated risk factors, etiology, and pathophysiology of cancers that occur less frequently. There is a gap in the current literature on the association between cancer development and SLE, and it is not known if there is an increased incidence of cancer development in persons with SLE.

Problem Statement

Chronic non-communicable diseases (CNCDs), such as SLE, account for a significant number of deaths, which are often a result of the impact of SLE on various organ systems (Manzi, 2009). However, the increased control of SLE disease activity has allowed researchers to study other areas of concern relative to persons with SLE (Manzi, 2009). In this study, I assessed the association between cancer and SLE.

Various types of cancer (including NHLs, cervical cancer, and lung cancer) are more prevalent in patients with SLE as compared to persons without SLE (Bernatsky et al., 2005; Kiss et al., 2010; Parikh-Patel et al., 2008). Researchers have proposed several risk factors that may lead to higher rates of cancer in persons with SLE, such as the use of immunosuppressive agents, which has been determined to be a carcinogen (Bernatsky et

al., 2006). Other risk factors that have been suggested, although not fully explored, include genetic predisposition such as the presence of antiphospholipid antibodies (which are frequently present in persons with both SLE and cancer), lifestyle-related risk factors such as smoking, and abnormalities in cell death regulation (Kiss et al., 2010; Tincani et al., 2010; Schultz & Harrington, 2003). However, it is uncertain whether SLE is associated with a higher rate of cancer (Bernatsky et al., 2005).

Previous research in assessing cancer development in patients with SLE has been based upon small study samples and cohorts that were not closed, which could increase the number of patients lost to follow-up. These studies lacked definitive diagnosis dates for SLE in the participants (Bernatsky et al., 2005; Parikh-Patel et al., 2008). As a result, previous studies were underpowered, and the scholars were unable to determine conclusively whether a relationship exists between SLE and cancer (Bernatsky et al., 2005). Moreover, the smaller sample sizes did not permit those diagnosed at a lower incidence to be assessed because there are many risk factors associated with cancer development associated with cancers that do not occur at a high frequency. This study used a large, population-based database to test the relationship between SLE and cancer. The larger sample size in this study allowed cancers that occur at a lower incidence to be examined. Finally, I examined the incidence rate of cervical cancer in patients with SLE.

Purpose of the Study

A propensity, score-matched, retrospective cohort study among SLE and non-SLE female patients identified in the CPRD was used to assess the association between cancer development and SLE. Female SLE patients with prevalent and/or incident cases were

included in the study. Non-SLE patients were matched with SLE patients in CPRD. The patients were then linked to an additional database for information on covariates. The cohort design was most suitable for this study to take advantage of all available data in the secondary data sources, to allow multiple outcomes to be examined, causality to be assessed, and to allow for the calculation of disease rates in the exposed and unexposed patients (Song & Chung, 2010).

The main exposure was SLE and non-SLE (unexposed), and study patients were designated as SLE or non-SLE. Both prevalent and incident cases were included in the analysis. The outcome was overall incidence of cancer and incident of cervical cancer. Cancer diagnoses were identified in CPRD (Health & Social Care Information Center. 2014). Incident cases included cases in which cancer was first diagnosed after the index date in people with at least 12 months of registration in CPRD. The first diagnosis of each cancer type was used in instances where the person had multiple cancer diagnoses at different times.

Data on covariates came from CPRD and HES. Information relating to SLE flares, number of hospitalizations, treatment, diagnoses, and medications came from HES. Each SLE patient was matched to non-SLE patients using variables to determine the probability of developing SLE (p [SLE]) in the patients. Balance between the exposed (SLE) and non-exposed (non-SLE) cohorts was obtained by using variables that were associated with the development of SLE to determine the treatment and non-treatment cohorts. This method of cohort selection was used to balance the SLE causing variables

in the cohorts. The control of the number and similarity of variables in the patients in each cohort kept the variables balanced between the study participants.

Research Questions and Hypotheses

RQ1: Is the risk of cancer development increased in SLE patients compared to non-SLE patients?

$H_0$1: There is no increased risk of cancer development in SLE patients compared to non-SLE patients for overall cancer risk and cervical cancer risk.

H_a1: There is an increased risk of cancer development in SLE patients compared to non-SLE patients for overall cancer risk and cervical cancer risk.

Theoretical Framework

The disease causation theory was the conceptual framework for this study. In this study, I assessed the association between cancer and SLE as compared to rates of cancer in non-SLE patients. There are many challenges in determining associations, especially in chronic diseases such as SLE and cancer. These two diseases involve multiple factors that interact and result in their disease state. The complexity of these two CNCDs led to the selection of a multifactorial causation theory as the basis for this study. The basic components of the multifactorial framework are that diseases have many causes, which cannot be independently attributed as the sole causative agent (Broadbent, 2009). The multifactorial framework was relevant for this study because of the multiple agents that have been proposed to cause SLE and to cause cancer, and because disease activity in SLE can contribute to an increased risk of cancer development.

The disease causation theory has been used to explain CNCDs such as diabetes, cancers, and autoimmune diseases (Najman, 1980). Unlike communicable diseases that can be attributed to a microorganism, these diseases do not originate from an organism, nor are they transmitted communicably. These factors of origination and method of transmission disqualify these diseases from being explained by the mono-causal model of disease that was applicable to communicable infectious diseases prevalent in the 19th century (Najman, 1980). Whereas diseases, such as smallpox, could be traced to a causative organism, diseases such as SLE and cancer cannot be traced to a single causative organism; these diseases are proposed to be caused by multiple factors such as environment, genetics, lifestyle choices, and possible organisms that result in the disease state (Broadbent, 2009). These multiple factors may explain the development of these diseases although no single factor by itself leads to the development of cancer or autoimmune diseases. The multifactorial disease causation theory was used as the foundation of this study because both SLE and cancers are the result of the occurrence of multiple factors.

Nature of the Study

A matched, retrospective, cohort study among SLE and non-SLE patients identified in the CPRD was used to assess the association of cancer development in patients with SLE. SLE patients with prevalent and/or incident cases were included in the study. Non-SLE patients were matched with SLE patients (4 to 1) on age, gender, practice, and date of registration in the CPRD. The patients were linked to the death registry for death due to cancer and to the HES database for data on some covariates (see

Table 1). The cohort design is most suitable for this study to take advantage of all available data in the secondary data sources, the CPRD, and the death registry. The cohort design allowed for multiple outcomes to be examined, causality to be assessed, and rare cancer exposures to be investigated. This design allowed for the calculation of disease rates in the exposed and unexposed patients (Song & Chung, 2010).

Table 1

Definitions

Variable	Definition	How Measured
Age (independent)	Number of years a patient has been alive	Numerical
Female (independent)	Male or Female	Categorical (M=0; F=1)
Immunocompromising Treatment (independent)	Medications that cause the body's immune system to be compromised	Numerical
Active Disease (independent)	Diagnosis of SLE as per American College of Rheumatology (ACR) or No diagnosis of SLE has been made	Categorical (Severe or Mild)
Cancer Development (dependent)	Diagnosis of a cancer	Categorical

The methodology used was quantitative, and the design was a high-dimensional propensity model. The main exposure was SLE (exposed), and study patients who did not have SLE were designated as SLE or non-SLE. Both prevalent and incident cases were included in the analysis. The outcome was cancer overall and cervical cancer in particular. Malignant cancer diagnoses were identified in CPRD, HES, and the death registry (for cause of death due to cancer). Incident cases included cases in which cancer

was first diagnosed after the index date in people with at least 12 months of registration in CPRD. The first diagnosis of each cancer type was used in instances where the person had multiple cancer diagnoses at different times.

Disease severity was measured by the use of IC treatment, along with other treatments such as NSAID and antimalarial drugs, which can be a proxy for disease activity and was considered a time dependent variable. A level of disease severity (e.g., mild, severe) was assigned based upon the number of comorbidities, number of hospitalizations, severe complications (e.g., renal transplantation), and use of IC treatments. The definition of this variable depended on what was available in the database.

Assumptions

Assumptions critical to the meaningfulness of this study that could not be demonstrated to be true included that data entered into the CPRD were accurate. The CPRD has been noted as having a high rate of data entry by the practitioners and was used in several studies for this reason. It was also assumed that patients diagnosed with SLE were accurately diagnosed, and the non-SLE patient did not have SLE.

Scope and Delimitations

An increase in life expectancy for persons with SLE has resulted in the recognition of other chronic diseases as cause for death in patients with SLE. Previous studies of cancer development in SLE patients have speculated that medications commonly used to treat SLE may be contributing factors. This study assessed whether the risk of cancer development was increased in SLE patients compared to non-SLE patients.

I chose to focus on overall cancer development and cervical cancer in particular. The cohorts consist of women only, which is the reason that cervical cancer was focused upon in this study. I specifically chose to not focus on cancers that have a higher incidence in the male population because only women will be included in this study. By focusing on overall cancer development and cervical cancer development in women, the findings are generalizable to women with SLE.

Limitations

The validity of this study depended on the accuracy and reliability of the information entered into the database. SLE is a difficult diagnosis to make, and there may have been a risk of including non-SLE patients. I excluded cases that were diagnosed within 6 months because both SLE and some cancers may have similar symptoms, and it may not be clear which diagnosis preceded the other. This scenario may threaten the validity of the study by possibly including patients who were inaccurately diagnosed with SLE or who had not been diagnosed with SLE when they should have been.

In a clinical study, bias refers to any errors in the study (Gerhard, 2008). The identification of bias or potential bias in a study allows the researcher to include steps that can assist in counteracting areas of bias. The most common areas of bias occurred in the history of the patients admitted into the study, the inclusion criterion used to select patients into the study, and the methods used to analyze the study results (Gerhard, 2008).

The study population included females with SLE and matched female non-SLE patients. Patients with a diagnosis of SLE were matched to non-SLE females using propensity scores that are based on SLE risk factors. This method of matching the

patients by using risk factors associated with SLE was one method to address bias that can enter a study by using the history of the patients in the comparison groups. Risk factors associated with the development of SLE were used to select the participants in both comparison groups. This matching decreased bias due to the selection process. A lack of consideration of the history of the patients, especially those factors that could lead to the development of SLE, could create bias between the comparison groups. The propensity score method of matching also addressed bias that is commonly a result of inclusion and exclusion criterion for a study. Selection using propensity score matching ensured that the comparison groups were equal in their covariates that could lead to the development of SLE and cancer.

Limitations are inevitable in research. CPRD is a combined effort of the MHRA's General Practice Research Database (GPRD) and the Department of Health's NIHR Research Capability Program (RCP) into a secure electronic research database. The use of data that have already been collected has limitations in that the research questions must be tailored around the information that is available in the database. An additional limitation associated with using CPRD is that I am not sure if the data were entered into the database with accuracy and reliability. CPRD does have the ability to be linked to other major databases such as the United Kingdom and Ireland Association of Cancer Registries (UKIACR) at a cost. The lack of funds to have CPRD linked with the UKIACR created another limitation in that the source of the most comprehensive data on cancer was not used in the study. The HES database was used because the UKIACR database could not be linked in this study.

Significance

Better control of SLE flares has resulted in fewer deaths caused by disease activity in this population of patients. Chronic diseases such as cancers have now been shown to be causes of death in patients with SLE. Current cancer screening recommendations for SLE patients may not be enforced enough to encourage patients to adhere to them. It is necessary to understand the association of cancer development in patients with SLE to create positive social change, which will support adherence of cancer screening recommendations.

Previous studies on the association of cancer development in patients with SLE have shown variable outcomes. Because several of the previous studies have been small and the comparison groups did not seem to be balanced as far as exposure to risk factors associated with SLE and cancer development, this study included a large population and balanced the cohorts based upon risk factors of SLE and cancer. The larger size of this study resulted in results that are generalizable. The study design, a high dimensional propensity model, allowed the patients to be balanced based on their exposure to medications and other diseases. This balance is significant because of the many factors that have been found in previous studies to contribute to chronic diseases such as cancer.

Summary

As discussed in this chapter, persons diagnosed with SLE are being diagnosed earlier, which has resulted in the use of medications that decrease the devastation cause to body organs because of flares. These factors have resulted in persons with SLE living longer than they did 25 years ago. This increase in lifespan has allowed researchers to

find that patients with SLE are now having chronic diseases such as cancer to cause mortality in SLE patients. Several studies have been conducted to assess the association of cancer development in patients with SLE, but many of the studies had a small study population and the results have had variability in the findings. My study used a large population-based database, the CPRD, to assess that association of cancer development in SLE patients. The large study population and the use of a high dimensional propensity model to select the cohorts resulted in cohorts that were balanced in terms of medications used and concurrent diagnoses experienced in patients in each of the cohorts. The balance of the cohorts in this study, because of the use of a high dimensional propensity model, allowed an equitable assessment of cancer development in patients with SLE.

In Chapter 2, a synopsis of the current literature on the problem will be presented with the strategy used to research the literature and a literature review of key elements related to concepts, variables, and methodology. A description of how the present study fills the gaps in the literature and how the results of this study will extend knowledge in the discipline will also be presented.

Chapter 2: Literature Review

Introduction

CNCDs such as SLE account for a significant number of deaths (Manzi, 2009). Persons with SLE are 2.4 times more likely to die of any cause than a non- SLE person after adjusting for demographic characteristics (Bernatsky, 2006). SLE is difficult to diagnose because it involves multiple organ systems, and it has no single diagnostic marker; rather, several different clinical symptoms and laboratory values must be combined eventually to diagnose the disease (Gill, Quisel, Rocca, & Walters, 2003). Diagnosis of SLE is made after the patient has exhibited 4 of 11 clinical symptoms and/or laboratory criteria. The average person with an autoimmune disease has a lifespan shortened by 15 years, and autoimmune diseases are the eighth leading cause of death among females (Nakazawa, 2008). The difficulty is diagnosing SLE and the impact of the lifespan has a significant economic impact of the public health system.

As previously stated, the economic burden caused by autoimmune diseases makes them a public health concern. The health care burden is estimated at approximately $120 billion, compared to $70 billion for cancer (Nakazawa, 2008). Additional burdens are placed on the health care system because of multiple trips to health care professionals in an attempt get an accurate diagnosis. In addition, patients with autoimmune diseases often face a poor quality of life that includes physical changes in their appearance (i.e., hair loss, facial rashes, loss of job, and/or eventual disability status) (Bertsia, Cervera, & Boumpas, 2012). The control of SLE flares has decreased the damage on organ systems in the body, which has resulted in decreased deaths due to SLE activity; this increased

control of SLE disease activity has now allowed researchers to study other areas of concern relative to persons with SLE (Manzi, 2009). In this study, I assessed the association of malignancies in persons with SLE.

Various types of cancer, including NHLs, cervical cancer, and lung cancer, increase in patients with SLE as compared to persons without SLE (Bernatsky et al., 2005; Kiss et al., 2010; Parikh-Patel et al., 2008). Interestingly, people with SLE were also found to be less likely to die of cancer (except for NHL and lung cancer) as compared to non-SLE persons (Bernatsky et al., 2006). Persons with SLE had a SMR of 2.8 for NHL and a SMR of 2.3 for lung cancer (Bernatsky et al., 2006).

Researchers have proposed several risk factors for the development of cancer among persons with SLE that would support an increased risk of cancer in this population, including the use of immunosuppressive agents, which increases the development of malignancies (Bernatsky et al., 2006). Other risk factors that have been suggested, although not fully explored, include genetic predisposition such as the presence of anti-phospholipid antibodies, which are frequently present in persons with both SLE and cancer; lifestyle-related risk factors such as smoking; and abnormalities in cell death regulation (Kiss et al., 2010; Schultz & Harrington, 2003; Tincani et al., 2010). However, whether or not SLE is associated with a higher rate of cancer is uncertain (Bernatsky et al., 2005). Previous research in this area has been based upon small study samples and cohorts that were not closed, which could increase the number of patients lost to follow-up, and the studies have lacked definitive diagnosis dates for SLE in the participants (Bernatsky et al., 2005; Parikh-Patel et al., 2008). As a result, previous

studies were generally underpowered to determine conclusively whether a relationship exists between SLE and cancer (Bernatsky et al., 2005). Moreover, the smaller sample sizes did not permit an examination of cancers that are diagnosed at a lower incidence, which is important because of the various associated risk factors, etiology, and pathophysiology of cancers that occur less frequently. This study used a large, population-based database to test the relationship between SLE and cancer in a large population. The larger sample size in this study allowed cancers that occur at a lower incidence to be examined. Finally, this study examined the incidence rates of various cancer types found in patients with SLE.

In this chapter, a discussion of literature on the pathogenesis of SLE and of cancer will be reviewed, followed by highlights of etiological similarities between the two chronic diseases. In addition, other autoimmune diseases that have established associations to specific cancers will be highlighted. I will then provide a review of literature written and based upon completed studies on the association between cancer and SLE with a focus on the types of cancers that have been found to have an increase in incidence in those studies. Finally, a review of literature on the medications used to treat SLE will be discussed.

Literature Search Strategy

The literature review should consist of studies that are similar to the research topic (Creswell, 2009). The literature contained in this review was used to establish a relationship with my study to research that has already been conducted. The principle source for obtaining the sources for the literature review was the MedImmune Corporate

Library. The MedImmune Corporate Library allows access to multiple electronic resources such as e-journals (greater than 12,000), online books (greater than 7,000), and dozens of databases such as Biosis, Embase, Medline/PubMed, SciFinder, Scopus, and Web of Science. The Walden University Library EBSCO databases were also used for literature searches. Other databases that were accessed included the Cumulative Index to Nursing & Allied Health Literature (CINAHL), MEDLINE, and ProQuest. Governmental Web sites, such as the U.S. Centers for Disease Control and Prevention (CDC) and the Medicines and Healthcare Products Regulatory Agencies (MHRA), were also used.

Strategies used when searching for and identifying articles for this literature review included the use of key words to search for articles (see Table 2). Key words and word combinations used to conduct searches included *systemic lupus erythematosus, autoimmune diseases, cancer and autoimmune diseases, malignancies, cancer development,* and *autoimmunity.* Once I found the articles, I reviewed the abstracts and read through the articles to determine if they would make a contribution to my understanding of the research topic. I also used guides to terms to locate articles (Creswell, 2009). Using multiple resources and databases allowed me to find several literature articles that were useful in this research project.

Table 2

Search Terms and Combinations for Literature

Topic(s)	Key words	Combinations
Theory	theories	chronic disease theories causation theory
SLE etiology	systemic lupus erythematosus SLE Lupus	lupus & autoimmunity lupus pathogenesis
cancer etiology	cancer malignancy lymphomas	cancer pathogenesis
common links incidence of cancer cancer development		cancer & SLE lupus & cancer lymphomas in lupus auto antibodies in cancer and lupus
drugs and cancer	lupus treatments	cancer in lupus

Articles published within 10 years of data collection were included in the study, unless there was limited information or published research on the topic or the research had a significant impact on the scientific body of knowledge of cancer or autoimmunity. I retrieved approximately 200 articles of which approximately 100 were found to be related to my study. Several articles were eliminated because the date of publication was greater than 10 years. Other articles were eliminated because the results appeared to be biased, or the journals were not peer-reviewed.

Theoretical Foundation

Because both SLE and cancer have a plethora of proposed causes, the disease causation theory served as framework for this study. The disease causation theory

proposes that diseases have many causes, which cannot be independently attributed as a sole causative agent (Broadbent, 2009). This multifactorial framework served as a basis to assess the use of immunosuppressive therapies and cancer development in persons with SLE. This framework also served to assess the development of specific types of cancers, such as cervical cancer and lung cancer in persons with SLE.

The disease causation theory was used to build the conceptual framework for this study. This study assessed the association of malignancy in patients with SLE. There are many challenges in determining associations especially in chronic diseases such as cancer and SLE. These two diseases involve multiple factors that interact and result in their disease state. The complexity of these two CNCDs resulted in the selection of a multifactorial causation theory as the basis of this study. The basic components of the multifactorial framework are that diseases have many causes, which cannot be independently attributed as the sole causative agent (Broadbent, 2009). The multifactorial framework was most relevant for this study because of the use of immune compromising treatments for SLE treatment and disease activity in SLE can both contribute to increased risk of cancer development.

The disease causation theory has been applied to explain CNCDs such as diabetes, cancers, and autoimmune diseases. Unlike communicable diseases that can be attributed to a specific microorganism, these diseases do not originate from an organism, nor are they transmitted communicably. These factors of origination and method of transmission disqualify these diseases from being explained by the mono-causal model of disease that was very applicable to communicable infectious diseases that were so

prevalent in the nineteenth century. Whereas diseases such as smallpox could be traced to a causative organism, diseases such as SLE and cancer cannot be traced to a single causative organism, but are proposed to be caused by multiple factors such as environment, genetics, lifestyle choices, and possible organisms that result in the disease state. These multiple factors together may explain the development of these diseases although no single factor by itself leads to the development of cancer or several of the autoimmune diseases. The multifactorial disease causation theory was used as the foundation of this study because both SLE and cancers are the result of the occurrence of multiple factors.

Etiology of SLE

SLE is a chronic inflammatory autoimmune disorder that occurs when the body's immune system attacks its own tissues and organs. An overactive immune response by the body against substances and tissues that are in the body, along with an inability to tolerate self-antigens, characterize the development of autoimmune diseases (Cristaldi, Malaguarnera, Rando, & Malaguarnera, 2011). Nagy, Koncz, and Perl (2005) described the etiology of SLE as a decrease in tolerance to self-antigens with polyclonal activation of B lymphocytes, production of different autoantibodies, and alteration in the function of T cells. This alteration in the T cells impacts T cell homeostasis and the modulation of immune responses to allergens, cancer cells, and pathogens (Belkaid, Piccirillo, Mendez, Shevach, & Sacks, 2002). SLE could be summarized as an immune response against internal nuclear antigens, which have been released from cells that were programmed to die but have been reactivated because they were presented to T cells. The reactivation of

these cells results in the production of helping B cells, which produce autoantibodies and secrete cytokines (regulatory proteins) that interact with cells of the immune system to mediate the immune response (Bertsias et al., 2012).

Etiology of Cancer

Cancer is a broad term that describes more than 100 diseases in which cell division gets out of control and invades other tissues in the body by spreading via the blood and/or lymphatic systems (National Cancer Institute [NCI], 2014). The body is composed of cells that grow, divide, and die; they are then replaced with new cells. However, cancer occurs when the cells genetic material gets damaged and the cells stop the normal cycle of dying, and their growth gets out of control (NCI, 2014). Damaged genetic materials of cells produce mutations, which affect normal cell growth (Bertsias et al., 2012). Cancers can either spread to other areas of the body (metastatic), or it can be contained in just one area of the body (benign). In 1999, cancer replaced heart disease as the leading cause of death among men and women 85 years of age and younger (Siegel, Ward, Brawley, & Jemal, 2011).

Many cancers have been associated with activities, elements in the environment, medications, and with disease processes. Age increases the likelihood of developing some types of cancer because of the longer exposure to potential carcinogens in the environment, foods, and other factors (Extermann, 2000). The risk of developing certain cancers in patients with rheumatic disease varies and has been found to have a higher prevalence depending upon the rheumatic disease (see Table 3). Several drugs and classes of drugs, including immunosuppressive agents, have been associated with cancer

development (Extermann, 2000). (Drugs used to treat autoimmune diseases will be discussed in depth later in this paper.)

Table 3

Rheumatic Diseases and Associated Malignancies

Rheumatic Disease	Organs Primarily Impacted by the Rheumatic Disease	Most Common Types of Cancer	Cancer Origin
Rheumatoid Arthritis	Joints	Lymphomas	Blood & lymph
Primary Sjögren's Disease	Mucous membranes	Lymphomas	Blood & lymph
Primary Scleroderma	Integumentary	Alveolar cell carcinoma	Respiratory
		Non-melanoma skin cancer	Integumentary
		Adenocarcinoma of the esophagus	Gastrointestinal (GI)
SLE	Various	Lymphomas	Blood & lymph
Celiac Disease	GI	Hematological cancers	
Inflammatory Bowel Disease	GI	Gastrointestinal cancers	

Turesson and Matteson (2013) noted that persons with rheumatoid arthritis have been found to have an increased risk of developing lymphomas, which are cancers that originate in the immune system and more specifically the lymphatic system. This finding is significant because several research studies have found HL and NHL to be elevated in patients with SLE. Both of the cancers originate in the immune system and more specifically in the lymphatic system.

Malignant lymphomas are classified according to the cell of origin and the biological understanding of the cell type; they have been indexed by the World Health Organization (WHO) (Word & Matasar, 2012). The malignant lymphomas are heterogeneous and originate from either B cells or T cells (Word & Matasar, 2012). B lymphocytes (B cells) fight against viruses and bacteria by producing proteins called antibodies, which attach to the germ so that other cells in the immune system know that they need to destroy them (American College of Rheumatology [ACR], 2013). T cells either destroy the marked germs, or they release other substances, which will digest the marked germs (ACR, 2013). The WHO has classified more than 50 types of lymphomas. However, I focused primarily on two types: non-Hodgkin's (NHL) and Hodgkin's lymphomas (HL). In 2011, approximately 66, 360 cases of NHL cases and approximately 8,830 cases of HL cases were diagnosed (Siegel et al., 2011). Previous studies have shown that NHLs and HLs have both been elevated in patients with SLE.

HL results from the malignant transformation of a B cell at either the post-germinal or post-germinal center stage of development (Word & Matasar, 2012). Of the 8,830 cases of HL diagnosed in 2011, approximately 1,300 resulted in death (Siegel et al., 2011). Diagnosis is more prominent in patients aged 20 to 29 years or greater than age 50 years; in addition, HL has a higher incidence in males than in females (Word & Matasar, 2012). In approximately 40% of HL cases, Epstein-Barr virus is detectable, and it is thought to be directly involved in the transformation to cancerous cells (Word & Matasar, 2012). This association of a virus in the transformation to cancerous cells has

also been a possible precursor in the transformation of the body producing antibodies against itself as in autoimmune diseases.

NHL also starts in the lymphocytes. There are several subtypes of NHLs but I primarily focused on the diffuse large B cell lymphoma (DLBCL). Mucosa-associated lymphoid tissue (MALT) type NHLs are often found following an autoimmune or chronic infectious process (Word & Matasar, 2012). The pathophysiology involves chronic stimulation of B lymphocytes by a persistent infection or autoimmune phenomenon that results in the cells cloning themselves (Word & Matasar, 2012). The DLBCL may occur as a transformation of a previously existing slow lymphoma that may or may not have been previously diagnosed (Word & Matasar, 2012). A weakened immune system and certain types of infections contribute to the development of NHLs. Certain infections have also been attributed to the production of autoantibodies, which act against the body, as in autoimmune diseases.

A greater incidence of lymphoproliferative cancers in autoimmune diseases should be expected because the chronic activation of B cells and T cells that occur in autoimmune diseases could serve as a catalyst to cancer development (Turesson & Matteson, 2013). Ragnarsson, Grondal, and Steinsson (2003) determined that patients with SLE are at increased risk for cancer development; surprisingly, prostate cancer was found to be increased in men with SLE. There was speculation that the increased prostate cancer rate could be possibly attributed to etiological mechanisms in males with an autoimmune disease, such as increased levels of antibodies against estrogen receptors,

which result in a decrease in the protective effect of estrogen for prostate cancer (Ragnarsson et al., 2003).

Common Links between Cancers and Systemic Lupus Erythematosus

Researchers have proposed several risk factors for the development of cancer among persons with SLE that would support an increased risk of cancer in this population, including the use of immunosuppressive agents, which increases the development of malignancies (Bernatsky et al., 2006). Other risk factors that have been suggested, although not fully explored, include genetic predisposition such as the presence of anti-phospholipid antibodies, which are frequently present in persons with both SLE and cancer; lifestyle-related risk factors such as smoking; and abnormalities in cell death regulation (Kiss et al., 2010; Schultz & Harrington, 2003; Tincani et al., 2009).

Both SLE and cancers are the result of cells in the body behaving in a manner that was not intended. These abnormal cell activities have some commonality in that (a) they have patterns of dysregulation that are similar such as autoantibodies in the blood; (b) they have bidirectional linkages as evidenced by clinical manifestations; and (c) immunosuppressive drugs used to treat these diseases can have an impact in cancer development (Achenza & Selmi, 2012). All cancers originate in cells of the body, which are the basic units of life (NCI, 2014). The fact that both cancers and autoimmune diseases originate at the cellular level is a primary link that may be instrumental in understanding whether the two disorders have an association with each other. The overstimulation of B cells along with a defective immune system contributes to the

greater propensity for lymphomas to be found in persons with autoimmune diseases (Cristaldi et al., 2011).

Abu-Shakra, Ehrenfeld, and Shoenfeld (2002) proposed that both autoimmune diseases and cancers have common etiologic agents such as environmental factors, the use of immunosuppressive agents, genetic factors that render them susceptible, and disturbances in the immune system. In the case of cancer and autoimmune diseases, it is difficult to determine which began first and whether the advent of one created a suitable environment for the other to begin. In addition, whether the use of treatment drugs caused the development of either lymphoma or an autoimmune disease is also questionable. These are the major issues faced when assessing the association of cancer in autoimmune diseases. Because of the heterogeneity of both cancers and autoimmune disorders, a causal relationship can be difficult to determine with certainty. Rosenquist (2008) found associations between inflammation, infectious agents, and certain lymphomas. The genetic make-up of individuals plays a role in the development of SLE and cancer as does the diet; the air quality; exposure to drugs, other than ones used for treatment for one of the disorders; and habits, such as smoking and tanning, which significantly impact the association. These are just a few of the numerous factors that may affect the association of cancer and SLE. These factors can cause a great impact on an association between cancer and autoimmune diseases.

Bei et al. (2009) studied the interaction between autoantibodies in cancer patients versus autoimmune patients. The authors found an overlap of antibodies that must be considered because they can change the properties of each other and can also affect the

growth and progression of each disease. Most studies conducted to date have been relatively small in size and unable adequately to examine the development and exposures of cancers that are not common. Outcomes from this study revealed that there is an association of cancer development in patients with SLE. More specific association of cervical cancers was also assessed in this study. The study outcomes could promote positive social change by reinforcing current recommendations for cancer screenings particularly in persons with SLE. Stronger adherence to cancer screening recommendations could enhance the ability to detect a cancer early enough so that treatment can be implemented that may result in a higher likelihood to effectively eradicate the cancer and decrease mortality due to cancer in persons with SLE. Results from this study may also equip persons with SLE with scientifically-based knowledge that will enable them to make decisions regarding their care with a clearer understanding of the cancer risks inherent to persons with SLE, particularly when factored with their knowledge of their personal familial risks for cancer development.

Incidence of Cancer in Autoimmune Diseases

Various types of cancer, including but not limited to cervical cancer and lung cancer, are thought to be increased in patients with SLE when compared to persons without SLE (Bernatsky et al., 2005; Kiss et al., 2010; Parikh-Patel et al., 2008). Researchers have proposed several risk factors for the development of cancer among persons with SLE that would support an increased risk of cancer in this population, including the use of immunosuppressive agents, which increase the development of malignancies (Bernatsky et al., 2006). Other risk factors that have been suggested,

although not fully explored, include genetic predisposition such as the presence of anti-phospholipid antibodies, which are frequently present in persons with both SLE and cancer, lifestyle-related risk factors such as smoking, and abnormalities in cell death regulation (Kiss et al., 2010; Schultz & Harrington, 2003; Tincani et al., 2010).

Whether SLE is associated with a higher rate of cancer is uncertain (Bernatsky et al., 2005). Previous research in this area has been based upon small study samples, cohorts that were not closed, which could increase the number of patients lost to follow-up; moreover, the studies have lacked definitive diagnosis dates for SLE in the participants (Bernatsky et al., 2005; Parikh-Patel et al., 2008). As a result, previous studies were generally underpowered to determine conclusively whether a relationship exists between SLE and cancer (Bernatsky et al., 2005). The smaller sample sizes did not permit an examination of cancers that are diagnosed at a lower incidence, which is important because of the various associated risk factors, etiology, and pathophysiology of cancers that occur less frequently.

Cancer Development and Systemic Lupus Erythematosus Disease

Studies completed to date have shown variable relationships between cancer developments in persons with SLE. The positive associations between hematologic cancers and several autoimmune diseases have been repeatedly demonstrated in several studies to date. Bernatsky et al. (2008) studied immunosuppressive therapy in SLE patients and found that immunosuppressive therapy may not be the principle driving factor for overall cancer risk, but may contribute to an increased risk of hematological cancers and is likely a plethora of causes that result in cancer development in SLE

patients. Broadbent (2009), in an examination of disease causation models, confirmed that a thorough understanding of the effects of various factors that contribute to a disease is necessary to measure an association of the variable to disease development. Kiss et al. (2010) identified several malignancies such as NHL, cervical cancer, and bronchial carcinomas, which are found in patients with SLE, are increased, and SLE patients have an increased incidence and risk of cancer development with the highest risk occurring in the first year of disease diagnosis. Nived et al. (2001) followed SLE patients to determine the rate of new malignancies following initial SLE diagnosis. Parikh-Patel et al. (2008) determined that patients with SLE have an increased risk of developing hematologic, kidney, and thyroid cancers. Hildalgo-Conde et al. (2013) actually suggested the incidence of cancer was four times greater than expected in a cohort study of Spanish patients.

Systemic Lupus Erythematosus Treatment Drugs and Cancer Development

Over the years, many drug classes have been found to be carcinogenic (Azab et al., 2008). Unlike many autoimmune diseases, SLE has only recently had a drug approved for the treatment of the disease. Until belimumab (Benlysta™) was approved for the treatment of SLE, other drugs were used as off-label treatments. The standard of care for SLE includes anti-inflammatory drugs, corticosteroids, antimalarial drugs, immunosuppressive drugs, and anticoagulants (Azab et al., 2008). Anti-inflammatory drugs, such as aspirin, acetaminophen, and non-steroidal anti-inflammatory drugs (NSAIDs) such as ibuprofen, are used to decrease inflammation. Corticosteroids are used to decrease swelling, warmth, tenderness, and pain associated with inflammation.

Antimalarial drugs are used for skin rashes. Immunosuppressive drugs are used to control inflammation when corticosteroid drugs have not effectively controlled the inflammation. Anticoagulants are used in SLE patients to prevent blood clots that can result from decreased mobility in people with SLE.

Bernatsky et al. (2008) proposed that immunosuppressive medications used to treat SLE may not be the main factor for overall cancer risk in SLE patients. Immunosuppressive medications may contribute to an increased risk of developing hematologic cancers, and suggests that future studies should evaluate other factors that increase the risk for malignancy in patients with SLE. Chang et al. (2005) found that use of NSAIDs to treat inflammation increased the risk of NHL. A study by Engels et al. (2005) found Sjögren's syndrome to have the strongest association in a study to assess immune-related conditions and immune-modulating medications as risk factors for NHL.

The theoretical framework for this study was the disease causation theory. The basic component of the disease causation theory is that diseases are caused by multiple factors and cannot be independently attributed to a sole causative agent (Broadbent, 2009). The multifactor framework was most useful because of the use of immune-compromising treatments for SLE treatment and disease activity in SLE, both of which could possibly contribute to increased risk of cancer development. Determining the cause for CNCDs is challenging because they involve multiple factors that interact and result in the disease state. This study involved two CNCDs, SLE, and cancers. Complexities of these two CNCDs required the use of a multifactorial causation theory.

Summary and Conclusions

Studies completed to date have shown variable relationships between cancer developments in persons with SLE. The smaller sample sizes of previous studies have not allowed an examination of cancers that are diagnosed at a lower incidence, which is important because of the various associated risk factors, etiology, and pathophysiology of cancers that occur less frequently. There is a gap in the current literature available on the association of cancer development in SLE patients, and it is not definitively know if there is an increased incidence of cancer development in SLE patients.

In this study, careful consideration was given to the multiple factors that have an interaction in SLE and in cancer. A large, population-based database was used to assess the relationship between SLE and cancer in a large population. The larger sample size in this study allowed cancers that occur less frequently to be examined. Finally, this study examined the incidence rates of various cancer types found in patients with SLE. The design of the study and the methods used to assess the findings were dynamic enough to allow for a detailed analysis that can be easily understood. The next chapter will provide details of the research design and methodology that was used in this study. A discussion of threats to validity will also be included in the upcoming chapter.

Chapter 3: Research Method

Introduction

In Chapter 2, I presented the results of several research studies on the possible risk factors for the development of cancer among persons with SLE. They included the use of immunosuppressive agents, genetic predisposition (presence of antiphospholipid antibodies), lifestyle-related risk factors such as smoking, and abnormalities in cell death regulation (Bernatsky et al., 2006; Kiss et al., 2010; Schultz & Harrington, 2003; Tincani et al., 2010). However, whether SLE is associated with a higher rate of cancer is uncertain. A notable limitation with the existing research in this area (small study samples, low retention rates, and a lack of a definitive and consistent diagnosis of SLE in the study participants) renders uncertainty about the ability to apply rates found to the general population (Bernatsky et al., 2005; Parikh-Patel et al., 2008). Moreover, the small sample sizes of several of the studies presented in the previous chapter prohibit an examination of cancers that are diagnosed less frequently. Less frequently diagnosed cancers are important because of the various associated risk factors, etiology, and pathophysiology of tumors that may occur less frequently.

This study was designed to address the limitations of previous research. I used a population-based database (CPRD) with a large sample size, which permitted an examination of cancers that occur at a lower incidence. Follow-up was not problematic because all data had already been collected and archived in the respective database. SLE was defined as any patient who had been diagnosed with SLE as per the American College of Rheumatology criteria for diagnosis of SLE.

In this chapter, the study design will be described along with the justification for the use of the analytical approach including its advantages over other models and limitations. A description of the target population will also be presented. The calculation of the minimum sample size needed to show an adequate statistical power to detect a treatment effect will be provided, along with the parameters and assumptions used to perform this computation. Finally, the procedure for gaining access to data archived in the CPRD will be described along with a description of anticipated threats to validity and potential ethical issues that were encountered during the study.

Research Design and Rationale

A matched, retrospective, cohort study among SLE and non-SLE patients identified in the CPRD was conducted to assess the association of cancer development in patients with SLE. SLE patients, including both prevalent and incident cases, were used in the study. Incident cases were defined as those with the first diagnosis of SLE after the last date of current registration (CRD) and the practice up-to-standard (UTS) date plus 12 months. The CRD was the date that the patient's information was entered into the CPRD; whereas, the practice UTS date was the date that the CPRD staff confirmed that the data met their internal quality standards. Prevalent cases were defined as those that had the first SLE diagnosis prior to the later of the CRD date and the UTS date plus 12 months. The index date for incident cases was the date of SLE diagnosis. The index date for the prevalent cases was the later of the CRD date and UTS date plus 12 months. The comparison group was a non-SLE cohort selected from CPRD. Non-SLE patients must not have had a SLE diagnosis at any time during the period considered for the SLE

cohort. This exclusion criterion prevented patients who experienced drug-induced lupus erythematosus (DILE) from entering the study. DILE can result after taking certain drugs and usually resolves on its own. Because DILE has a different pathophysiology than SLE, it was not included in this study.

Randomization that occurs in a randomized controlled trial (RCT) theoretically ensures a balance of measured and unmeasured covariates amongst the experimental and the control groups (Austin, 2011). However, in an observational study, the patients are assigned to a group, but the groups will differ in systematic ways. This method of group assignment did not take into consideration the fact that patients in either group may have had concurrent diseases, may have been undergoing treatments, and may have possessed other characteristics that could interfere with an equitable evaluation of the outcome (Sugihara, 2010). Group assignment without a method to control for confounding characteristics introduces selection bias for the study because the study results were not generalizable because of the numerous confounding variables for both cancer and SLE.

Studies of association of cancer development in patients with SLE have not used a propensity score methodology until this study. The study design allows the cohorts to be balanced better than previous studies that have assessed cancer development in patients with SLE. Sugihara (2010) recommended the use of propensity score methodology in observational studies in which multiple confounding characteristics are present in the patients. This method helps to balance the differences between comparison groups. The propensity score is the predicted probability given a determined set of measured covariates (Rassen, Glynn, Rothman, Setoguchi, & Schneeweiss, 2012). The propensity

score allows an observational (nonrandomized) study to be designed and analyzed, so that it mimics many of the characteristics of an RCT (Austin, 2011). This method is based upon the creation of a similar distribution of observed baseline covariates between the two groups, and it was used in this study to help to balance the measured differences between the comparison groups.

The propensity score analysis in this study began by identifying risk factors for SLE. A logistic regression model was used to determine the risk of developing SLE. The predicted probability of developing SLE p(SLE) was going to be calculated by the regression of the covariates on a dichotomous variable for whether a patient has SLE. This calculation would have produced a probability of SLE for each woman in the cohort. Patients with and without SLE would have be matched based on the p(SLE) (Caliendo & Kopeinig, 2005). This matching could only occur within the area of common support, which is the range of p(SLE) that is common to both the SLE and non-SLE women, so it would exclude women with very high and low probability of having SLE.

Once the risk factors that contribute to the development of SLE were determined, a test to check if the matching procedure balanced the distribution of the relevant variables in the control and treatment groups would have been performed (Caliendo & Kopeinig, 2005). A D-statistic test would have been used to determine if the SLE and non-SLE patients had significant differences. The purpose of this step would have been to balance the cohorts, somewhat analogous to randomization in an RCT (Caliendo & Kopeinig, 2005).

After determining the two cohorts, the risk factors for cancer development were used to determine the predicted probability of patients in the SLE cohort and in the non-SLE cohort to develop the outcome of cancer p(CA). This was going to be done by again using a logistic regression model. Next, the relative risk (RR) would have been determined using proportional hazards regression (Cox, 1972). The exponentiated coefficients on each covariate are an odds ratio of the odds of cancer. A survival analysis would have been done using the proportional hazard regression model.

The use of a high dimensional propensity model allowed the computer program to search the database and determine common variables in patients with SLE. In this case, I used concurrent diagnoses and medications. Austin (2011) described propensity scores as using baseline characteristics to assign patients to a treatment group. In this study, the average treatment effect (ATE) was the outcome framework that was used. The ATE is the average effect at the population level when the whole group is moved from untreated to treated (Austin, 2011). The other framework that could have been used was the average treatment effect for the treated (ATT). The ATT is the average effect of treatment only for patients who actually received treatment (Austin, 2011). The ATE framework was more advantageous in this study because it was more important to estimate the effect of cancer development in persons with SLE versus in patients without SLE (Austin, 2011).

Because an observational database was used, the comparative effectiveness of covariates was estimated by using the propensity score. This computer-assigned propensity score was used to compare individuals with similarly estimated scores by

either matching or stratification (Curtis, Hammill, Eisenstein, Kramer, & Anstrom, 2007). Matching by propensity scores would have matched patients in one treatment group directly to a patient in another treatment group solely based on the propensity score. To use stratification, the difference in the ATE of the two groups and then the average effect would have been calculated within each stratum (Curtis et al., 2007). The number of potential patients for inclusion in the study and the use of stratification would have increased the probability similarities in the two comparison groups; however, it may have been difficult to distinguish the patients from one another (Curtis et al., 2007). Inverse Propensity Weighting (IPW) would employ less distributional assumptions about the data, prevent additional confounding, could incorporate time dependent covariates, and could also manage censored data (Curtis et al., 2007).

In addition to using propensity scoring to match patients for participation in an observational study, the treatment effect could also have been assessed by evaluating the multiple variables identified (Heinze & Jüni, 2011). Use of propensity scoring for observed covariates for cohort selection in this study would have assisted in decreasing the many sources of bias, which result from SLE. Both of these diseases have suggestive causes and overlapping factors that may predispose a person to the development of either or both diseases. In this study, matching of SLE patients to non-SLE patients using conventional randomization measures would have introduced bias into the analyses because of the numerous risk factors associated with the outcome of cancer development (Heinze & Jüni, 2011).

After matching the patients by propensity scores, they were going to be linked to the HES database for data on some covariates for reasons described later in the archival data section. The cohort design was most suitable for this study to take advantage of all available data in the secondary data source CPRD. The cohort design allowed for a thorough examination of those persons who developed cancer (Song & Chung, 2010). The longitudinal data contained in CPRD was also used in the cohort design as opposed to a case control design.

Probability of Systemic Lupus Erythematosus

The primary exposure was SLE, and study patients were designated as SLE or non-SLE. To calculate the probability of being in the SLE group, I considered the risk factors for disease in the propensity score approaches. SLE is most prevalent in people of color (e.g.., African, Asian, Indian, or Hispanic) (CDC, 2014). The onset of SLE usually occurs between the ages of 16 to 44 years (CDC, 2014). Several exogenous and endogenous risk factors contribute to the development of SLE. In a case control study conducted in Sweden to explore risk factors associated with the development of SLE, a history of hypertension had an odds ratio (OR) of 3.7 (1.4-9.8), history of a reaction to the sun had an OR of 2.3 (1.1-4.8), history of a drug allergy had an OR of 3.6 (1.4-9.5), and a family history of SLE had an OR of 6.8 (1.4-32) (Bengtsson, Rylander, Hagmar, Nived, & Sturfelt, 2002). The presence of these risk factors independently may not result in the development of SLE, but a combination of these factors, along with other factors, may increase the chances of developing SLE. The high OR for a family history of an autoimmune disease aligns with the known genetic predisposition that is associated with

many autoimmune diseases. Cooper, Dooley, Treadwell, St. Clair, and Gilkeson (2002) had similar findings in The Carolina Lupus Study, which was conducted in several counties in North and South Carolina. SLE risk factors that were to be used to match the SLE and non-SLE cohorts are listed in Table 4.

Table 4

SLE Risk Factors

Risk Factor	Definition	How Measured	Study-Relevant Associated SLE
Year of Birth (proxy for age)	Year of birth	Continuous	All
Race	Race or ethnicity reported by the patient	African/Black=0 Hispanic=1 Asian=2 Indian=3	All
Hypertension	Diagnosis of hypertension at any time	Yes or No	All
Ultraviolet Skin Reactions	History of skin reactions	Yes or No	All
Family History of SLE or Other Autoimmune Diseases	Documentation of self-reported family history of any autoimmune disease	Yes or No	All
History of Allergy to Antibiotics	Diagnosis of a history of an allergy to an antibiotic	Yes or No	All
History of Shingles or Cold Sores	Self-report or diagnosis of shingles or cold sores	Yes or No	All

The original plan was to use a logistic regression model to determine the risk of developing SLE. The predicted probability of developing SLE p(SLE) was going to be

determined by matching the covariates between the SLE and non-SLE women in CPRD. This calculation would have been used to identify the cohorts of SLE and non-SLE patients by matching the risk factors for SLE in both populations. The matched patients who fall within the area of common support would have made up the SLE and non-SLE cohorts in the study.

Probability of Cancer

The outcomes for this study included incident cancer (all types) and cervical cancer in particular. Cancer diagnoses were identified in the CPRD and HES. Incident cases included cases in which cancer was first diagnosed after the index date in people with at least 12 months of registration in the CPRD. The first diagnosis of each cancer type was used in instances where a person had multiple cancer diagnoses at different times.

Numerous risk factors may increase the chance of developing cancers (WHO, 2014). All of the cancer causing risk factors cannot be addressed in one study because of the limitations using the existing data in the CPRD and HES database. In this study, therefore, I analyzed some of the most common risk factors that are routinely included in the medical history of a database, such as the CPRD. The cancer risk factors in this study were to include age, tobacco use, human papillomavirus (HPV) infection, Epstein-Barr virus (EPV) infection, overweight, pregnancy history, and oral contraceptive (OC) use.

Age (greater than 65 years) is a risk factor for most types of cancers (CDC, 2014). As a person ages, he or she is exposed to a multitude of elements, including environmental and lifestyle factors that are associated with an increased risk of general

cancer development (CDC, 2014). The CDC determined that environmental exposures, such as air pollution, secondhand tobacco smoke, asbestos, drinking water containing large amounts of arsenic, and pesticides, are linked to some cancers such as lung, skin, and bladder. Lifestyle elements, such as cigarette smoking, tobacco use, infections, radiation, immunosuppressive medicines, diet, alcohol, physical activity, and obesity, are either known to increase cancer risk or may affect the likelihood of cancer (CDC, 2014).

Exposure to tobacco, such as smoking, snuff, and chewing tobacco, is the most preventable risk factor for lung cancer, and it contributes to cancers such as mouth, nose, throat, larynx, esophagus, liver, bladder, kidney, pancreas, colon, rectum, cervix, stomach, blood, and bone marrow (CDC, 2014). Smoking renders a person exposed to many cancer-causing chemicals that affect multiple body organs when the chemicals are carried via the blood system to the organs. The chemicals act to damage the DNA of cells and may contribute to the development of multiple cancers (American Cancer Society [ACS], 2014). A history of tobacco use was a cancer risk factor in this study and was measured as either yes or no.

A history of infection with certain viruses and some bacteria can increase the risk of developing cancer (CDC, 2014). Infections with certain viruses and bacteria increase the risk of cancer development because the infection changes the person's DNA, and changes the behavior of the cells which may cause them to replicate at a rate greater than necessary, which is cancer (Cancer Research United Kingdom, 2014). In this study, I analyzed the impact of EBV and HPV in cancer development of lymphoma and cervical

cancer respectively. HPV is a routine test that is a part of the Papanicolaou (PAP) test for females.

A person that is overweight or has a Body Mass Index (BMI), which is a measure of body fat based upon height and weight, has been attributed as a risk factor of developing cancer. Being overweight, which can be a result of a poor diet or a lack of adequate physical activity, increases the risk of cancers such as colon, uterus, prostate, esophagus, breast, and kidney (CDC, 2014). In addition to increased BMIs and increased levels of estrogens and insulin, which are also associated with being overweight or obese, are some of reasons for increased cancer development (Cancer Research United Kingdom, 2014.

A full-term pregnancy before the age of 17, as compared to a woman who had the first pregnancy after the age of 25, and women that had more than three pregnancies are associated with an increase in the risk of cervical cancer development (CDC, 2014). These two factors increase the risk of cervical cancer development because they increase the chances of acquiring a HPV infection due to the potential of the woman having unprotected sex with a greater number of sexual partners (ACS, 2014). The hormonal changes associated with pregnancy may also contribute to a weakened immune system and render the woman susceptible to HPV infection (ACS, 2014).

Use of estrogen-progestagen OCs over a 5-year or greater period may increase the development of cervical cancer (CDC, 2014). A study that analyzed the effect of OC use with a background of HPV infection showed that the use of an OC for more than 5 years by women that had a HPV resulted in increased rates of cervical cancer (Smith et al.,

2003). Although HPV exposure is known to be the most important cause of cervical cancer, when combined with long-term use of OCs, the relative risk (RR) of cervical cancer development was increased as duration of use increased (Smith et al., 2003). Smith et al. studied women with cervical cancer with no history of OC use (RR 1.9-2.2) compared to women who had used OCs (RR 1.6-3.9) for more than 10 years and found the incidence of cervical cancer increased with longer use of OCs.

Other risk factors could be included but would require a specially designed data collection tool. These other factors would include sunlight exposure; ionizing radiation exposure; exposure to chemicals such as asbestos, benzene, and cadmium; family history of cancer; alcohol use; and diet and physical activity. Because the CPRD does not contain the detailed risk factors mentioned, it was not possible to assess these cancer risk factors in this study. However, the cancer-associated risk factors that I had planned to be included in this study are listed in Table 5.

Table 5

Cancer Risk Factors

Risk Factor	Definition	How Measured	Study-Relevant Associated Cancer
Age (independent)	Number of years the patient has been alive at the time of CPRD entry	Numerical	All
Tobacco Use	Use of any tobacco product such as cigarettes, cigars, pipes, snuff, or chewing tobacco	Yes or No	Cervical All Cancers
HPV Infection	Positive HPV infection at any time	Yes or No	Cervical
EBV Infection	Positive EBV infection at any time	Yes or No	Cervical
Overweight	Having a body mass index (BMI) greater than 25	Yes or No	Cancer
Pregnancy History	Age of 1st pregnancy and number of pregnancies	Categorical	Cervical
Contraceptive History	Use of an OC	Yes or No	Cervical

Propensity scores were initially going to be used for matching, which would have balanced the differences between the patients included in the study. Propensity scoring was going to be used to evaluate the multiple variables that could be used to match the comparison groups (Heinze & Jüni, 2011). The decrease in the many sources of bias that result as related to SLE was to be controlled with the use of propensity scores.

Treatment effect was going to be assessed by evaluating the multiple cancer-related risk factors that were identified based on the literature review (Heinze & Jüni, 2011). Cancer and specifically cervical cancer have numerous strongly suggestive causes and overlapping factors that may predispose a person to development of either or both

diseases. None of the previous researchers performed an analysis of the outcome risk factors and incorporated this information into the overall measure of association.

The Database

Resource constraints encountered with the design choice include the cost to link data from the CPRD to other data sources such as the UK cancer registry. This limitation resulted in the reliance on HES data rather than data from the cancer registry. CPRD (2013) estimated that a link between the CPRD and the UK cancer registry would cost approximately $16,000. The cost would pay the CPRD staff to link data in the CPRD with data in the cancer registry. Limited funding did not allow me to establish the linkage to the UK cancer registry for this study because I was not able to secure the funds from my employer, nor was I able to personally support the costs to link the CPRD with the UK cancer registry. Because of this resource constraint, I decided to rely on data that could be found in the CPRD and HES.

Time constraints consistent with the design choice included working closely with programmers to obtain the proper data from the CPRD. Variables were defined as concepts, so that the needed variables could be retrieved from the CPRD. Once obtained, additional time was invested to review the data to confirm the inclusion and exclusion of variables that were not primary malignancies. In the initial plan, a method to match the patients in the other databases would have been implemented once the cohorts were selected. The matching would have taken a significant amount of time to ensure that the cohorts were matched appropriately. Once matched, several analyses and testing of the data would need to have been completed. There would have been a significant amount of

time required because of the design choice: however, the time constraint was not as high as it would have been if raw data had to be collected and assimilated, as in a primary data collection study. I recognized there were time constraints associated with this study, and allowances had to be factored in to allow the necessary time needed to complete this study successfully.

Frankfort-Nachmias and Nachmias (2008) explained that the non-interventional observational study design is made up of two groups for comparison: an exposed group and a control group. These two equal groups are either exposed to the independent variable (i.e., active disease) or not exposed to it (i.e., no active disease). In this study, the comparison groups included patients with SLE and non-SLE patients. The study observed whether the presence of SLE had an influence on cancer development (dependent variable). Equitable selection of the cohorts decreased bias, which would compromise associations detected in the study. Exposure to the treatment allowed for the evaluation of the treatment and an assessment to be made of how the treatment would affect the general population.

A quantitative study design was used in this study. A propensity score matched retrospective cohort study between a cohort of SLE and non-SLE patients identified in the CPRD database was initially planned for this study. The study assessed the association of cancer development in patients with SLE. SLE patients consisted of either prevalent or incident cases. Non-SLE patients were to be matched with SLE patients using the p(SLE). The patients were then to be linked to the UK Death Registry for death due to cancer. Data were linked to the HES database for additional information on the

covariates. The cohort design was most suitable because it allowed multiple outcomes to be examined, enabled causality to be assessed, was good for investigating rare exposures, and this design would also allow calculation of disease rates in the exposed and unexposed patients (Song & Chung, 2010).

This study assessed the association of cancer in patients with SLE. Because both SLE and cancer have many proposed causes, the disease causation theory served as framework for this study. The disease causation theory suggests that diseases have many causes that cannot be independently attributed to a sole causative agent (Broadbent, 2009). This multifactorial framework served as the basis to assess the identified cancer-related risk factors and cancer development in patients with SLE. This framework also helped to evaluate the development of particular types of cancers, such as cervical cancer in patients with SLE.

Methodology

Population

The source population was obtained from the CPRD. The study population included all females with SLE and with no SLE. All female SLE patients with a diagnosis of SLE included in the area of common support in CPRD (i.e., prevalent and incident cases) were included. Drug-induced lupus cases were excluded because they are of a different etiology than SLE. Patients with cutaneous lupus were excluded because they do not follow the same pathophysiology as SLE. The only restriction on age was that they be 18 years old. Patients were followed for up to 10 years. Because patients with a SLE

diagnosis prior to their registration in the CPRD, prior to their current registration date (CRD), were likely to get a cancer diagnosis entered within the first few months after registration, which results in the incidence of diseases being overestimated during this period (Lewis, Bilker, Weinstein, & Stron, 2005). Therefore, incident cases were defined as those with the first diagnosis of SLE after the latest of the CRD and the practice up-to-standard (UTS) date plus 12 months. Prevalent cases were defined as those that had the first SLE diagnosis prior to the later of the CRD date and the UTS date plus 12 months. The index date for incident cases was the date of SLE diagnosis. The index date for the prevalent case was the later of the CRD date and UTS date plus 12 months. The comparison group was the non-SLE cohort selected from CPRD. Patients must not have had an SLE diagnosis at any time during the period considered for the SLE cohort. Non-SLE patients were to be matched individually to SLE patients based upon the propensity weights.

Sampling and Sampling Procedures

All female SLE and non-SLE patients in the CPRD who fell within the area of common support were to be included in the study. An initial assessment of the CPRD database revealed that approximately 4,000 female patients in CPRD had SLE (both new and prevalent cases) and of those 4,000, roughly 2,000 could be linked to other registries. The necessary sample size was calculated using clincalc.com, which was provided by Clincalc, LLC. To determine the sample size needed adequately to power the study, the overall incidence of cancer in females in the United Kingdom used was 266/100,000 or 0.27% (Cancer Research United Kingdom, 2014). The incidence of cancer anticipated in

SLE women in CPRD used was 18.88/1000 or 1.89% (SAEfetyworks, 2015). The probability of type-1 error was 0.05 and the power was set at 0.80 with the enrollment ratio being 1 or equal cohort sizes (see Table 6). Using these numerical settings, it was calculated that the total study size needed to have sufficient statistical power to detect a difference (type II error) was 1,276 total with 638 patients assigned to each of the cohorts, SLE and non-SLE.

Table 6

Parameters Used for Sample Size Calculation

Incidence, group 1(non-SLE)	0.27%
Incidence, group 2 (SLE)	1.89%
Alpha	0.05
Beta	0.2
Power	0.8

Archival Data

The CPRD is funded by the National Health Services (NHS) National Institute for Health Research (NIHR) and Medicines and Healthcare products Regulatory Agency (MHRA). CPRD has been used in more than 890 clinical reviews and papers, and the database is considered the gold standard by many because the high compliance of health care providers entering the data into the database (Clinical Practice Research Database, 2013). The CPRD is managed by a group that serves the general practitioners who enter the data into the database, and the managing group serves researchers by anonymizing the data so that it can be linked to other databases (CPRD, 2013). The CPRD is a combined effort of the MHRA's General Practice Research Database (GPRD) and the Department

of Health's NIHR Research Capability Program (RCP) into a secure electronic research database. The NHS assigns a unique patient identifier number that is only used by a trusted third party for linkage, and it is never released to researchers (CPRD, 2013). This anonymized identification number ensures that linkages to other database are valid while maintaining privacy for the patient. Access and use of CPRD data are controlled under the laws of the United Kingdom and Europe.

Data in the CPRD are available online after completion of a 2-day training program provided by the CPRD research team. The training provides background information about CPRD, data fields contained in the database, linkages available, data entry information that practitioners follow, and services that CPRD can provide. There are costs for accessing the CPRD data to cover services provided by the CPRD Research Team. These costs are either paid by individuals, academic institutions, or business entities. My access to the CPRD was granted by my employer. The information obtained from this study will be useful as background information for patients with SLE. My employer conducts research to develop medicines, and findings from this study will be helpful in decisions related to studies being conducted for drugs that could potentially be used as a treatment for SLE. The company had already paid for a number of people that could have access to the CPRD, and I was approved to be one of the persons who could have access for the company; therefore, I did not have to pay out of my pockets for access to the CPRD. There are legal agreements that cover all aspects of the use of the CPRD data and services (CPRD, 2013).

There are linkages already established with the Office of National Statistics (ONS), which contains mortality data. These data are available to researchers after a research project is approved by the Independent Scientific Advisory Committee (ISAC) for MHRA database research (CPRD, 2013). The ISAC is an appointed expert advisory body that provides advice on research requests to access CPRD (CPRD, 2013). A request to ISAC must include the project's methodology, information on the medical, statistical, and epidemiological aspects of the proposed study (CPRD, 2013). ISAC can request additional information needed in making a determination whether the proposed study in CPRD will be approved or denied. All investigators and collaborators are included in the application to have the protocol approved by ISAC.

The CPRD is funded by the NHS, NIHR, and the MHRA (CPRD, 2013). The CPRD services are designed to improve the way NHS clinical data can be linked to enable many types of research and deliver useful research outputs (CPRD, 2013). Non-SLE patients were to be matched with SLE patients on propensity scores. A link to data in the HES database was made to obtain data on diagnosis of malignancies. The cohort design was most suitable for this study to take advantage of all available data in CPRD and HES database. Types and sources of information or data were CPRD provided demographic and diagnostic data on patients with SLE, and HES provided information about SLE diagnosis, cancer diagnoses, diagnostic and laboratory information.

Operationalization

HES is a data warehouse that contains details of approximately 125 million hospital admissions, outpatient appointments, and emergency records each year at

National Hospital Services hospitals in England (Health & Social Care Information Center [HSCIC], 2014). Data entered are administrative in that they allow hospitals to be paid for the care they deliver (HSCIC, 2014). Although administrative in nature, HES can be used to monitor trends and patterns in hospital activity, to assess efficient delivery of care, and to inform patient choice (HES, 2014). The HES database was developed in 1987 after a report that collected hospital information was commissioned by the NHS and published. The 1987 report was the start of a database that contained information compiled for each hospital admission and used to assess the severity of SLE disease activity for patients. Data in HES are collected monthly and are accessible via the Internet.

Data Analysis Plan

Research Questions

RQ1: Is the risk of cancer development increased in SLE patients compared to non-SLE patients?

Ho1: There is no increased risk of cancer development in SLE patients compared to non- SLE patients for overall cancer risk and cervical cancer risk.

Ha 1: There is an increased of cancer development in SLE patients compared to non-SLE patients for overall cancer risk and cervical cancer risk.

Planned Methodology: Conditional Logistic Regression

To measure incidence of cancer, patients were followed from the index date until the earliest of the first cancer diagnosis, death, loss to follow-up, or end of study. A

conditional logistic regression model would have allowed an analysis of binary outcome data with one or more predictors and where observations were not independent, but were matched (Statsdirect, 2014).

Planned Methodology: Proportional Hazard Regression

A survival analysis was initially planned, using the proportional hazard regression model with the index date for the SLE cohort being the earliest date indicated in CPRD presumed as being the onset date. The index date for the non-SLE cohort would be the date of enrollment in CPRD. The identified risk factors were to be analyzed to determine their impact on the development of cancer in patients with SLE. The Cox proportional hazard model (backward method) was initially planned to be used to examine the relative effect of each covariate on the incidence of cancer. All analyzes would be performed using Statistical Analysis System (SAS) version 9.2. SAS is software developed at North Carolina State University in 1976, and the software can be used for advanced analytics, data management, predictive analysis, and business intelligence (Statistics Solutions, 2014).

The original plan was to have risks for each cancer site to be determined by measuring the relative risk, which is the ratio of the probability of an event occurring in the exposed group to the likelihood of the event occurring in the comparison or unexposed group to the expected number of cancers. A cohort type study allowed calculation of incidence, which could be used to calculate relative risk by dividing the cumulative incidence in the exposed by the cumulative incidence in the unexposed. This measure would allow a comparison of the variables on the risk of cancer development.

The cohort design best used the benefits of using the longitudinal data contained in CPRD. All analyzes were performed using SAS version 9.2.

In multivariable analysis, consideration would be given to all cancer and cervical cancer known to be related to SLE as the outcome. The primary exposure was SLE or non-SLE. In addition to overall cancer development, a particular association of cervical cancer was included. Depending on the power, some cancers may need to be grouped together to increase power. Because I eventually assessed overall cancer and cervical cancer, it was not necessary to group cancers. Sensitivity analyses using different definitions for latency were going to be conducted to take latency of cancer into account (e.g., 1, 2, 3, 5 years). Although the size of the SLE cohort in CPRD is substantially larger than in many previous studies, an attempt was be made to take latency for cancer into account: however, this would be dependent upon the effective sample size and commensurate level of power.

Threats to Validity

The validity of this study depended on the accuracy and reliability of the information entered into the database. SLE is a difficult diagnosis to make, and there may be a risk of including non-SLE patients. As mentioned earlier, cases that were diagnosed within 6 months were to be excluded because both SLE and cancer have similar symptoms. They were excluded because it was impossible to determine which diagnosis occurred first. Allowing cases where it is unknown whether the SLE or cancer was diagnosed first would threaten the validity of the study.

Ethical Procedures

The IRB review and approval protected the study and the university by independently reviewing the methodology being used in the study to determine if there was any potential for harm to anyone participating in the study (Rudestam & Newton, 2007). This study underwent IRB review and was found that it did not pose an ethical concern because it was a retrospective study using secondary data from CPRD; the IRB approval number is 06-03-15-0149603. The study retrospectively assessed the association of cancer development in persons with SLE disease. The patients were linked to the HES database for additional data on the variables. All data were historical, and all patients remained anonymized during all reviews. All data will be stored on my computer at home and at my place of employment for 5 years after the study is complete. I have access to these data as well as my employer.

This study involved research to obtain information to assess a public health issue, cancer development in SLE patients. The study did not involve the provision or withholding of a medical intervention. This retrospective study did not have people participating in it; therefore, no determination of full disclosure of information to potential study participant was needed, which is an activity of the IRB.

Summary

It is uncertain if SLE is associated with a higher rate of cancer. Previous research in this area has been based upon small study samples; cohorts that were not closed, which could increase the number of patients lost to follow-up; and the studies have lacked definitive diagnosis dates for SLE in the participants (Bernatsky et al., 2005; Parikh-Patel

et al., 2008). Previous studies were underpowered; therefore, they did not allow a conclusive determination to be made about whether a relationship exists between SLE and cancer (Bernatsky et al., 2005). The smaller sample sizes were not conducive to an examination of cancers that occur less frequently. The examination of less frequent cancers is important because of the various associated risk factors, etiology, and pathophysiology of these cancers.

This study used a large, population-based database to test the relationship between SLE and cancer in a large population. The larger sample size in this study allowed examination of cancers that occur at a lower incidence. This study also examined the incidence rates of various cancer types found in patients with SLE. Non-SLE patients were initially going to be matched with SLE patients using IPWE. All participants in this study had to be at least 18 years of age and the study included a period of 10 years up to 2014. The patients were linked to the HES database so that data on other covariates could be collected. The cohort design allowed me to take advantage of all available data in CPRD and HES.

All SLE subjects in CPRD were included in the analysis. Overall incidence of cancer was the main outcome; in addition, the risk of developing cervical cancer was examined. Cancer diagnoses identified in the CPRD and HES were assessed to obtain additional information about the covariates. Year of birth, race, hypertension, ultraviolet skin reactions, family history of SLE or other autoimmune diseases, history of allergy to antibiotics, and history of shingles or cold sores were the covariates that I initially planned to analyze. The covariates that were planned to be included to determine the

probability of cancer were age, tobacco use, HPV infection, EBV infection, overweight, pregnancy history, and contraceptive history. Data on the covariates were obtained from the CPRD and HES.

The disease causation theory served as framework for this study. This theory acknowledges the understanding that SLE and cancer develop because of several factors together rather than any one single factor leading to the development of either of these diseases. This framework served to assess the development of specific types of cancers, such as cervical cancer, in persons with SLE. Statistical analyzes were performed using SAS.

In the next chapter, a review of the purpose of the study, research questions, and hypotheses will be presented. A detailed discussion of the descriptive and demographic characteristics of the cohorts will be given along with a correlation of the sample population with the population of interest. The results of all analyses performed will be presented and details of any challenges encountered with the implementation of the study. Descriptive characteristics and assumptions will be evaluated and all findings will be reported. Tables and figures will be displayed to illustrate the results as appropriate. A thorough examination of steps taken to execute the study will be presented along with a comparison to previous findings from similar research.

Chapter 4: Results

Introduction

Chronic non-communicable diseases such as SLE account for a significant number of deaths, which are often a result of the impact of SLE on various organ systems (Manzi, 2009). The increased control of SLE flares has decreased the devastation on organ systems in the body and has resulted in decreased deaths due to SLE activity. This increased control of SLE disease activity has now allowed researchers to study other areas of concern relative to persons with SLE (Manzi, 2009). The purpose of this study was to assess the association of cancer in patients with SLE. Because both SLE and cancer have a plethora of proposed causes, the disease causation theory was most appropriate to serve as framework for this study. According to the disease causation theory, diseases have many causes that cannot be independently attributed as a sole causative agent (Broadbent, 2009). This multifactorial framework served as a basis to assess the development of cervical cancer in patients with SLE

A high dimensional propensity weighted, retrospective cohort study among SLE and non-SLE female patients identified in the CPRD database was used to assess the association of cancer development in patients with SLE. Female SLE patients with prevalent and/or incident cases were included in the study. Non-SLE patients were matched with SLE patients in CPRD. The main exposure for this study was SLE (exposed) and non-SLE (unexposed) study patients, which was designated as SLE or non-SLE. The outcome was overall incidence of cancer. The first diagnosis of each

cancer type was used in instances where the person had multiple cancer diagnoses at different times.

Research Questions and Hypotheses

The research questions and hypotheses for this study were

RQ1: Is the risk of cancer development increased in SLE patients compared to non-SLE patients?

$H_0$1: There is no increased risk of cancer development in SLE patients compared to non-SLE patients for overall cancer risk as compared to non-SLE population and cervical cancer risk as compared to non-SLE population.

H_a 1: There is an increased risk of cancer development in SLE patients compared to non-SLE patients for overall cancer risk as compared to non-SLE population and cervical cancer risk as compared to non-SLE population;

In this chapter, the data collection details will be described. The plans for collecting the data were discussed in Chapter 3, and any deviations from those plans will be detailed in this chapter along with rationale for the deviation. The demographics of the patients will be presented, and an analysis of the patients' proportionality to the larger population will be discussed. The results of the analyses that justify the covariates that were used in the model will be provided and discussed. The statistical analyses that were used to determine and extract the sample populations will be presented along with the findings from the analyses that were performed. No additional statistical tests emerged from the analysis of the main hypotheses. Tables and figures will be included in this chapter to assist in illustrating the results as appropriate. Finally, the answers to the

research questions will be summarized, and the prescriptive materials that will be presented in Chapter 5 will be introduced.

Data Collection

The source population for this study was patients in CPRD. The study population included females with SLE. An inverse probability of treatment weights (IPTW) was performed because the estimate of interest was the ATE or the effect that would be seen if both the SLE and non-SLE cohorts received the same concurrent diagnoses and the same concomitant medications (Harder, Stuart, & Anthony. 2010). The IPTW allowed all of the study patients (SLE and non-SLE) to be weighted up to represent the entire study population (Harder et al., 2010).

All female SLE patients with a diagnosis of SLE (prevalent and incident cases) and all non-SLE females in the area of common support in CPRD were included in the study using propensity scores based on SLE risk factors. Patients had to be at least 18 years of age to be included in the study. Through CPRD, patients were followed for up to 15 years. The index date for this study was the date of SLE diagnosis. The comparison group was the non-SLE cohort selected from CPRD. Patients must not have had an SLE diagnosis at any time during the period considered to be included in the non-SLE cohort.

In Chapter 3, the data collection plan was developed based upon predefined covariates. These variables were selected from a review of published literature on risk factors associated with developing SLE and risk factors associated with the development of cancer. The predefined risk factors were to be used to first calculate the probability of developing SLE. Variables such as race, age, history of hypertension, history of a

reaction to the sun, and a family history of SLE or other autoimmune diseases were to be used as risk factors to match the SLE and non-SLE cohorts. Database limitations prevented me from using all of the risk factors. Later I will present the actual risk factors that were used to match the SLE and non-SLE cohorts.

Once the SLE and non-SLE cohorts were identified, the original plan was to determine the probability of cancer in both cohorts by using the same process. Risk factors for cancer development, as identified in literature reviews, were to be used to match the SLE and non-SLE cohorts. This would create a comparable population by controlling for prediction of SLE and the covariates that were chosen. Recognizing that all of the cancer causing risk factors could not be addressed in one study, some of the most common risk factors associated with developing cancer, such as increased age (greater than 65 years), tobacco use or exposure, HPV infection, EBV infection, being overweight, pregnancy history, and OC use, were to be used. Later, I will present the method actually used to assess cancer in the SLE and non-SLE cohorts.

Upon attempting to use the predefined risk factors to identify the SLE and non-SLE cohorts, I realized that CPRD did not identify race in the database. Age could be determined, but history of hypertension, history of a reaction to the sun, and a family history of SLE or other autoimmune diseases were not standard fields in CPRD. The lack of these data fields led me to re-evaluate the methods for selecting the cohorts for this study. A high dimensional propensity score model was used instead of the original study plan, which was to assign a weight to the predefined variables selected based upon the literature review.

In the high dimensional propensity model, a computer program was used to assign weights to variables based upon the analyses of the common variables related to SLE that were found in the database (Schneeweiss et al., 2009). In this study, I used the most common diagnoses and most commonly associated prescriptions used by persons with SLE. The selected covariates included the top 100 diagnoses and prescriptions that were found most frequently in the database for the SLE population. The covariates (concurrent diagnoses and concomitant medications) were then assimilated into a propensity score based confounder adjustment model (Schneeweiss et al., 2009). The diagnoses and medications were ranked amongst the SLE population in CPRD, and then they were matched with the non-SLE population in CPRD. The seven steps used to implement the high-dimensional substitute adjustment in CPRD are shown in Figure 1.

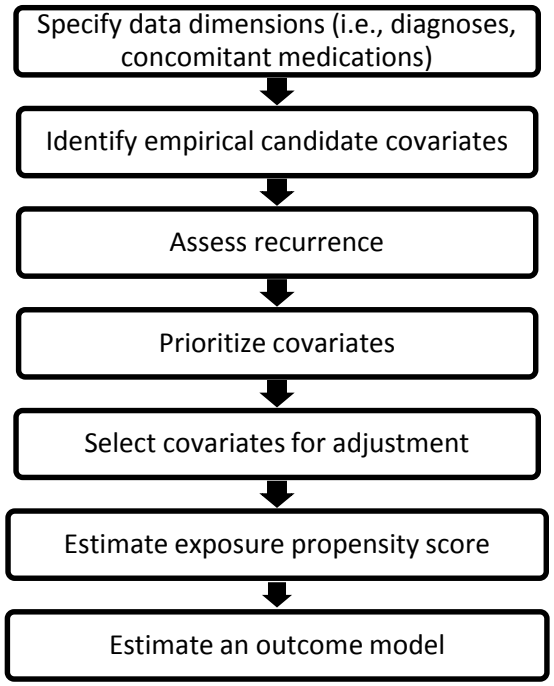

Figure 1. Flow chart for a basic high-dimensional propensity score algorithm
Note: (modified from Schneeweiss et al., 2009, Copyright © Lippincott Williams & Wilkins).

The CPRD is a health care use database, so data collected were not geared toward assessing clinical disease development and severity, such as the covariates that had been identified from the literature review. However, the CPRD does contain a large amount of data that can be used as proxy to describe disease status. In this study, the SLE and non-SLE cohorts were selected by using the concurrent diagnoses that were commonly diagnosed in patients with SLE, as proxy. After selecting the female patients with SLE in CPRD, the most common concurrent diagnoses that SLE patients had were selected. The concurrent diagnoses were used because SLE is usually accompanied by other disorders

because of the nature of the disease process of SLE and its impact on all systems of the body.

The top 100 diagnoses that were given to persons with SLE in the CPRD were selected. The diagnoses that were identified were ranked in the order of how often the diagnosis occurred in SLE patients. The concurrent diagnoses were described in groups of similar types of events. Of the 100 top concurrent type of diagnoses reported for patients with SLE in the CPRD, approximately 26% of the diagnoses were some form of pain, such as headache, general pain, or specific events of pain of various body areas. Pain is a common manifestation of SLE because the body produces antibodies against itself. The next largest group of events was infections and inflammations, which accounted for 15% of the concurrent diagnoses. Inflammation is usually an indication of some type of infection. In patients with SLE, inflammation can be a result of the breakdown of body processes because of SLE, or it could be a result of the enhanced development of infections because of using corticosteroids and other medications to treat the symptoms of SLE. The remainder of the concurrent events was found in the category of respiratory events (9%), rashes or skin disorders (8%), female-specific events (7%), gastrointestinal events (6%), malaise (5%), urinary events (5%), and cardiac events (3%) as displayed in Table 7. All of the concurrent diagnoses before the index date were descriptive of patients with SLE and the pathophysiology of the disease. A random date was selected among all encounters for the non-SLE population.

Table 7

Concurrent Diagnoses in SLE Patients

Event Category	
Pains/Aches	26%
Infections/Inflammations	15%
Respiratory	9%
Rash/Skin Irritations	8%
Female Events	7%
Gastrointestinal	6%
Malaise/Fatigue	5%
Urinary	5%
Cardiac	3%

The other proxy used to describe the SLE population was medications that the SLE patients in CPRD were prescribed prior to the index date. The 100 most frequently prescribed concomitant medications for SLE patients in CPRD were identified. The top 10 frequently prescribed concomitant medications with their drug class in parenthesis were acetaminophen (pain), amoxicillin (antibiotic), prednisolone (corticosteroid), codeine (pain), influenza virus vaccine (prophylactic), diclofenac (inflammation), hydroxychloroquine (antimalarial), ibuprofen (pain/inflammation), hydrocortisone (corticosteroid), and trimethoprim (an antibiotic) (see Table 8). All of the concomitant medications identified by the computer search were in alignment with the manifestations of the SLE disease process. Other concomitant medications identified were medications used as off-label treatments for symptoms of SLE.

Table 8

Frequently Prescribed Concomitant Medications in SLE Patients

Rank	Medication	Drug/Indication
1	Acetaminophen	Pain/Inflammation
2	Amoxicillin	Antibiotic
3	Prednisolone	Corticosteroid
4	Codeine	Pain

A medication specifically for the treatment of SLE was not available until 2013 when Benlysta™ was approved to be sold on the market. Up until 2013, medications prescribed for SLE patients were for palliative treatment of symptoms that the disease exhibits and not to cure the disease itself. As noted by the 100 most prescribed medications identified in the CPRD for the high dimensional propensity score modeling, none of the medications were specifically for the indication of SLE.

Pain and inflammation were the most commonly prescribed medications for patients with SLE in the CPRD. Pain and inflammation of the joints are common manifestations of SLE because the body produces antibodies that work against itself. Medications for pain and inflammation were the most commonly prescribed class of medication for patients with SLE in CPRD. Corticosteroids were the second most prevalently prescribed class of medication. Steroids are used to decrease swelling, warmth, tenderness, and pain that are associated with inflammation. Long-term steroid use can cause an increased risk of infections because it suppresses the immune system in

general. This immune system suppression secondary to corticosteroid use is the reason that antibiotics and anti-infectives are the third most frequently prescribed class of medication in the SLE population in CPRD.

The Sample

Once the risk factors that contribute to the development of SLE were determined and the cohorts were identified, an OR estimate was performed to determine if there was an increased probability of incurring SLE p(SLE) if the patient had been diagnosed with certain concurrent diseases. Another estimate was run to determine if there was an increased p(SLE), if the patient had been prescribed particular medications (Le, 2009). Of the 100 concurrent diagnoses made to SLE patients found in CPRD, the 10 diagnoses with the highest probability of being made in SLE patients are shown in Table 9. These findings are consistent with the earlier noted top 10 types of diagnoses found, with the greatest number of diagnoses being some type of pain, infection, and rash. Because the population studied was entirely female, female events are noted to be concurrently diagnosed often in the SLE patients identified in CPRD.

Table 9

Diagnoses with the Highest Probability of Being Made in SLE Patients

		95% Wald Confidence Limits	
Diagnosis	Point Estimate	Low	High
Herpes Zoster	27.428	21.059	35.723
Irritable Bowel Syndrome	16.431	13.63	19.808
Generalized aches and pains	1.906	1.6	2.27
Cardiac disease monitoring	1.877	1.541	2.285
Breast lump symptom	1.558	1.283	1.893
Vomiting	1.528	1.146	2.036
Arthritis	1.51	1.266	1.801
Chest pain	1.461	1.312	1.627
Hemorrhoids	1.461	1.213	1.76
Rash	1.455	1.307	1.619

Of the 100 concurrent medications prescribed to SLE patients found in CPRD, the10 medications with the highest probability of being prescribed to SLE patients are shown in Table 10. An antimalarial (hydroxychloroquine) was the highest prescribed, followed by a disease modifying antirheumatic drug (azathioprine), then a corticosteroid (prednisolone), followed by a blood thinner, antihypertensive, a NSAID, a thyroid hormone, a topical corticosteroid, another DSAID, and an iron supplement. These medication prescriptions are consistent with the pathophysiology of SLE.

Table 10

Medications with the Highest Probability of being Prescribed to SLE Patients

		95% Wald Confidence Limits	
Medication	Point Estimate	Low	High
Hydroxychloroquine	78.859	70.846	87.778
Azathioprine	4.589	3.885	5.42
Prednisolone	3.193	2.891	3.525
Warfarin	2.847	2.43	3.334
Nifedipine	2.026	1.735	2.365
Aspirin	1.579	1.4	1.781
Levothyroxine	1.443	1.249	1.668
Clobestasol	1.411	1.202	1.657
Diclofenac	1.362	1.245	1.489
Ferrous sulfate	1.303	1.171	1.45

It is important to note that propensity score matching is somewhat different from matching that occurs in a randomized study; propensity score matching only balances the observable whereas, in a randomized study both observable and unobservable distributions are balanced (Lee, 2013). Because the propensity score was used for the purpose of balancing, it was imperative to run a test to determine whether the propensity score effectively balanced the covariates used to identify the SLE and non-SLE cohorts. This balance statistic ensured that the propensity score had the same distribution for both the SLE group and the non-SLE group. I had to ensure that the treatment medications and the concurrent diagnoses found in CPRD were distributed in a balanced manner amongst the SLE and non-SLE cohorts.

An IPTW was performed because the estimate of interest was the ATE or the effect that would be seen if both the SLE and non-SLE cohorts received the same concurrent diagnoses and the same concomitant medications (Harder et al., 2010). As

demonstrated in Figure 2, the IPTW weighting technique assigns each individual patient a weight or propensity score that is the inverse probability of receiving the treatment that they actually received; this allows all of the study patients to be weighted up to represent the entire study population (Harder et al., 2010).

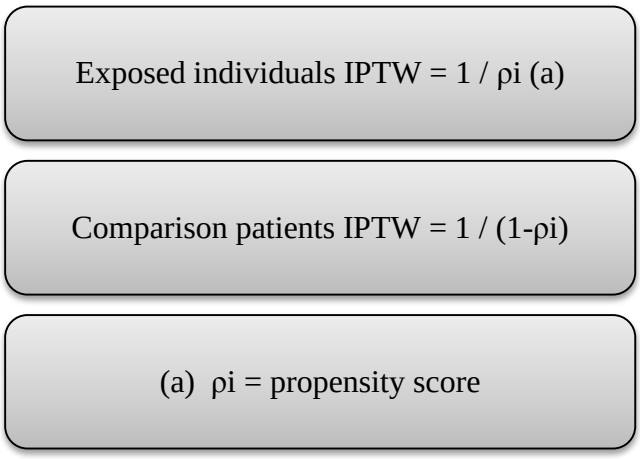

Figure 2. IPTW weighting technique to assign each individual patient a weight or propensity score

Heavily weighted covariates would create bias because they would have significantly more influence in determining the balance of the cohorts. Simply removing the covariates with the largest weights would create additional bias because the covariates with the largest weights are the best predictors of the outcome that is being compared. Therefore, a stabilization technique was applied to decrease the variability of the weights. The treatment and comparison weights were each independently multiplied by a constant that was equal to the expected value of being in either group (Harder et al., 2010). Once the weights were stabilized by the computer program, a technique known as trimming was used to minimize the influence of any remaining outlying weights.

Trimming limits the stabilized weight by shortening them to within a specific range. The goal was to select the trimmed percentile that is most aligned with the baseline. Figure 3 displays the balance of the diagnoses at baseline, with no trim, at 99% of SLE trimmed, and at 95% trimmed. A 95% trim was selected because that was most aligned with the baseline as depicted in Figure 4 with the no trim and 99% trim graphs removed so that the balance can be seen clearer. The 95% trimmed weights align most consistently with the baseline weight and therefore, a 95% trimming will be used for the weights of the diagnoses in CPRD.

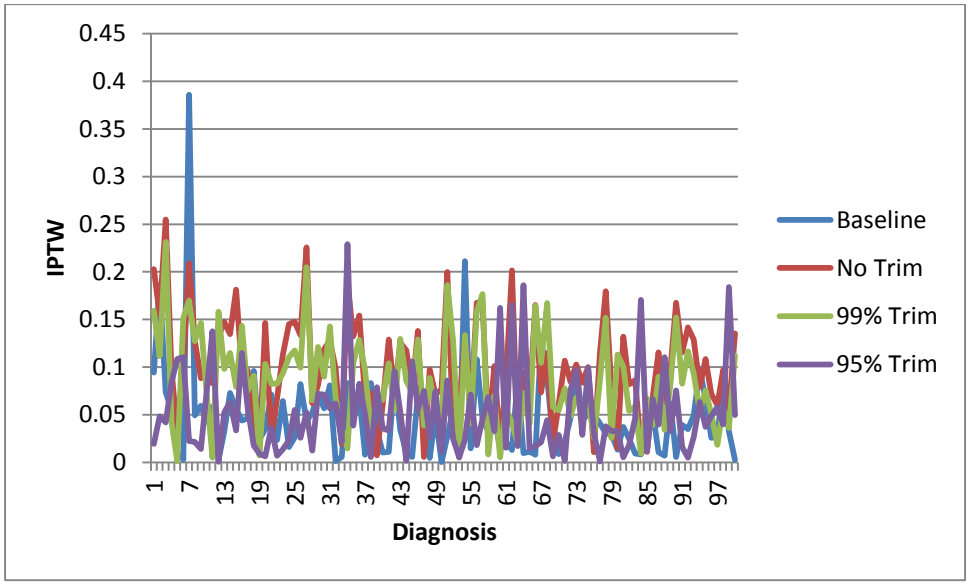

Figure 3. Balance of concurrent diagnoses using IPTW

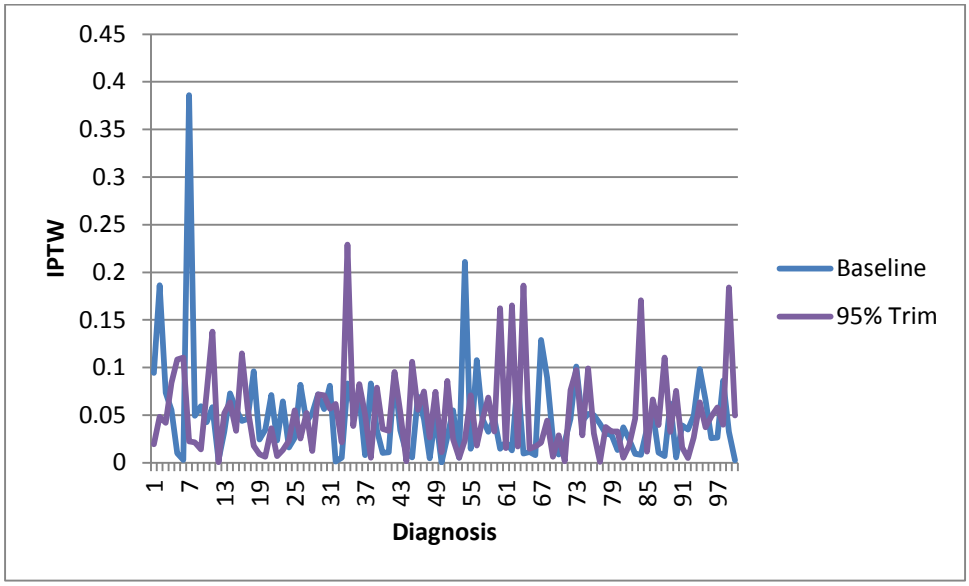

Figure 4. Balance of 95% trimmed weights for diagnoses using IPTW

Figure 5 and Figure 6 display the concomitant medications prescribed for SLE patients in CPRD with the balance of the IPTW weights at baseline, with no trimming, at 99% and at 95% trimmed.

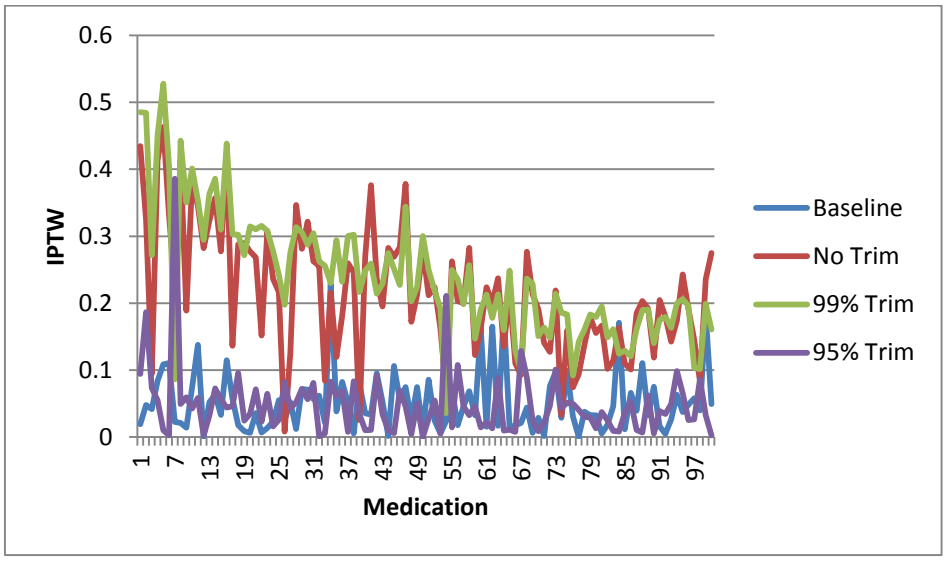

Figure 5. Balance of concomitant medications using IPTW

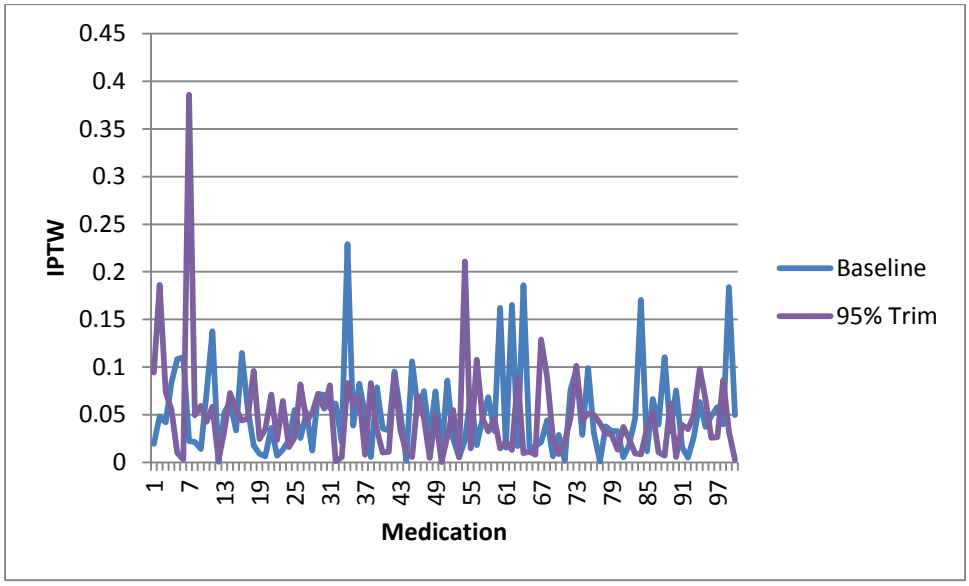

Figure 6. Balance of 95% trimmed weights for concomitant medications using IPTW

Using the concurrent diagnoses and concomitant medications for a high-dimensional propensity model rendered an initial SLE population of 3,362 and a non-SLE population of 2,719,084 in CPRD, using IPTW with a 95% trimming to best balance the weights of the covariates. The IPTW was then applied to the SLE and non-SLE cohorts that had been identified in CPRD. This application of the IPTW and the trimmed covariates rendered a SLE cohort of 3,025 and a non-SLE cohort of 180,555 patients.

The sample obtained is an accurate representative to the SLE and non-SLE populations in the general population. The SLE and non-SLE populations included in the sample are balanced in the concurrent diagnoses made to each population and they are balance in the concomitant medications prescribed to them. These balances make both cohorts equally exposed to covariates that may contribute to the development of SLE.

The equal exposure to potential covariates will allow a more equitable assessment of malignancy development amongst the two groups.

To identify primary cancers in CPRD, the Observational Medical Outcome Partnership (OMOP) vocabulary was used to identify the concept of cancer. The concept names of *neo, mal,* and *can* were used to identify any terms with any of those series of letters. The initial findings included a great number of events that were (a) benign events (curable by removal), (b) events that were potential precursors to malignancies, and (c) not malignancies. Because we only want to include malignancies in this study, and more specifically, the first malignancy diagnosed, we then added *pri* to the search. For this reason, secondary malignancy sites were not included nor were any of the previously mentioned conditions that were not malignancies. As to be expected, the greatest number of malignancies included female specific cancers such as breast, uterus, and cervix. The next largest group of malignancies was in the digestive system. The number of cervical cancers found in the non-SLE cohort was significantly greater than the number found in the SLE cohort. A depiction of the malignancies, of which there were at least 10 or greater events, found in the SLE or non-SLE cohorts in CPRD are listed in Figure 7 and to indicate that they belong to the same body system in Table 11 and Figure 8.

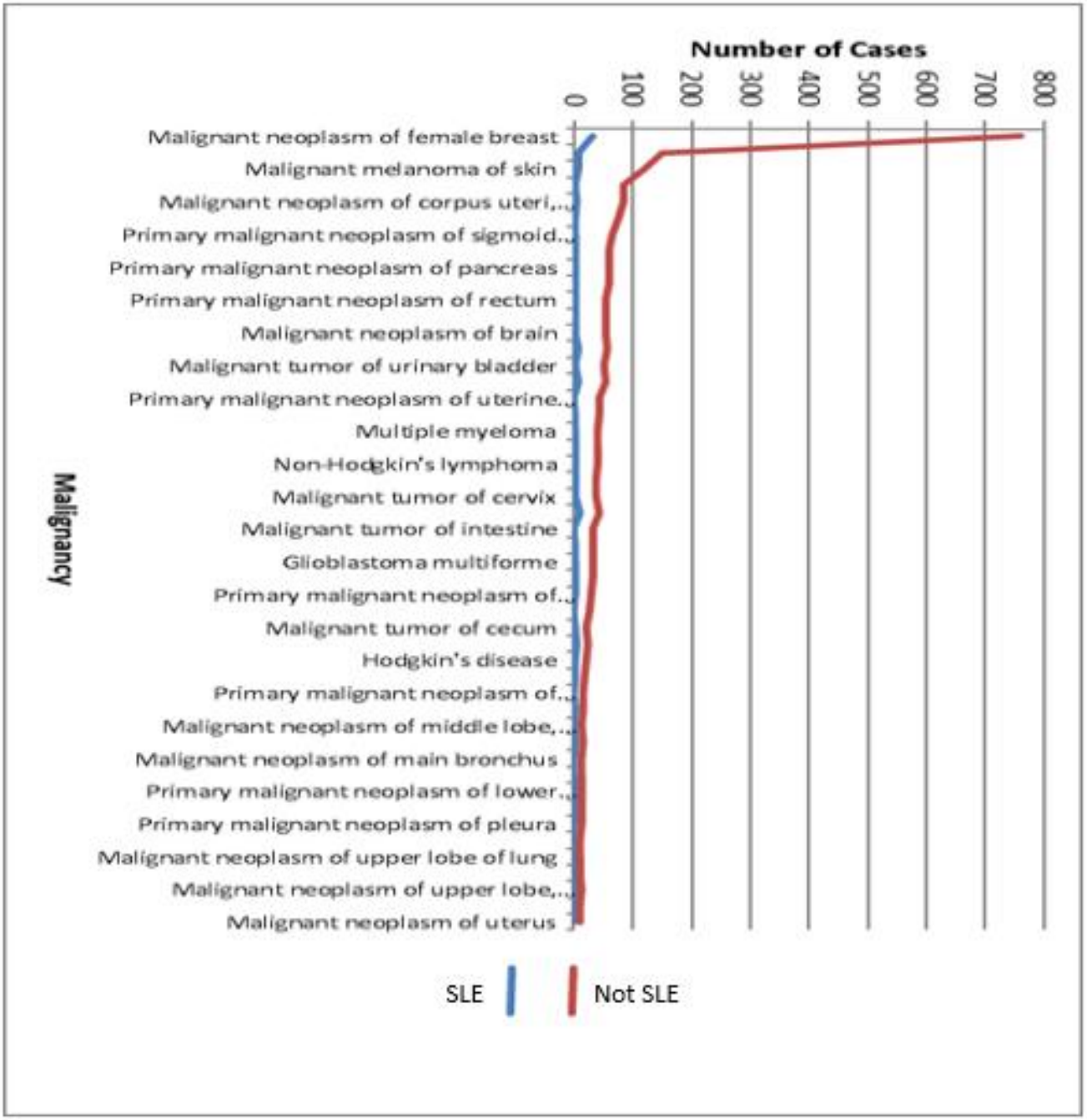

*Figure 7***.** Malignancies in SLE and non-SLE cohorts

Table 11

Malignancies by Body System

	SLE (*n*)	Non-SLE (*n*)
Female Reproductive Cancers		
Malignant neoplasm of female breast	32	728
Malignant neoplasm of corpus uteri, excluding isthmus	5	78
Primary malignant neoplasm of uterine cervix	0	42
Malignant neoplasm of endometrium of corpus uteri	1	36
Malignant tumor of cervix	1	36
Malignant neoplasm of body of uterus	0	11
Carcinoma of cervix	1	10
Primary malignant neoplasm of vagina	0	10
Malignant neoplasm of uterus	0	10
Digestive System		
Primary malignant neoplasm of colon	8	141
Primary malignant neoplasm of sigmoid colon	2	62
Malignant tumor of esophagus	2	57
Primary malignant neoplasm of rectum	1	54
Primary malignant neoplasm of esophagus	2	41
Primary malignant neoplasm of stomach	9	33
Malignant tumor of intestine	0	32
Primary malignant neoplasm of cecum	2	29
Primary malignant neoplasm of rectosigmoid junction	2	28
Adenocarcinoma of rectum	0	27
Malignant tumor of cecum	1	21
Malignant tumor of ascending colon	1	17
Primary malignant neoplasm of transverse colon	0	16
Primary malignant neoplasm of lower third of esophagus	0	14
Primary malignant neoplasm of anus	2	12
Carcinoma liver and/or biliary system	1	15
Primary malignant neoplasm of liver	1	14
Integumentary System		
Malignant melanoma of skin	7	113
Squamous cell carcinoma of skin	6	52
Epithelioma basal cell	0	14
Urinary System		
Primary malignant neoplasm of bladder	2	75
Malignant tumor of urinary bladder	1	50
Endocrine System		57
Primary malignant neoplasm of pancreas	3	56
Primary malignant neoplasm of ovary	3	53
Malignant tumor of ovary	2	37
Primary malignant neoplasm of thyroid gland	2	30
Primary malignant neoplasm of head of pancreas	1	

(table continues)

	SLE (*n*)	Non-SLE (*n*)
Nervous System		
Malignant neoplasm of brain	2	52
Glioblastoma multiforme	1	30
Cardiovascular System		
Myelodysplastic syndrome	8	46
Multiple myeloma	1	38
Lymphatic System		
Non-Hodgkin's lymphoma	3	37
Hodgkin's disease	1	19
Respiratory System		
Malignant neoplasm of middle lobe, bronchus or lung	0	14
Malignant neoplasm of main bronchus	0	14
Primary malignant neoplasm of pleura	0	12
Malignant neoplasm of upper lobe of lung	0	11
Malignant neoplasm of upper lobe, bronchus or lung	3	10

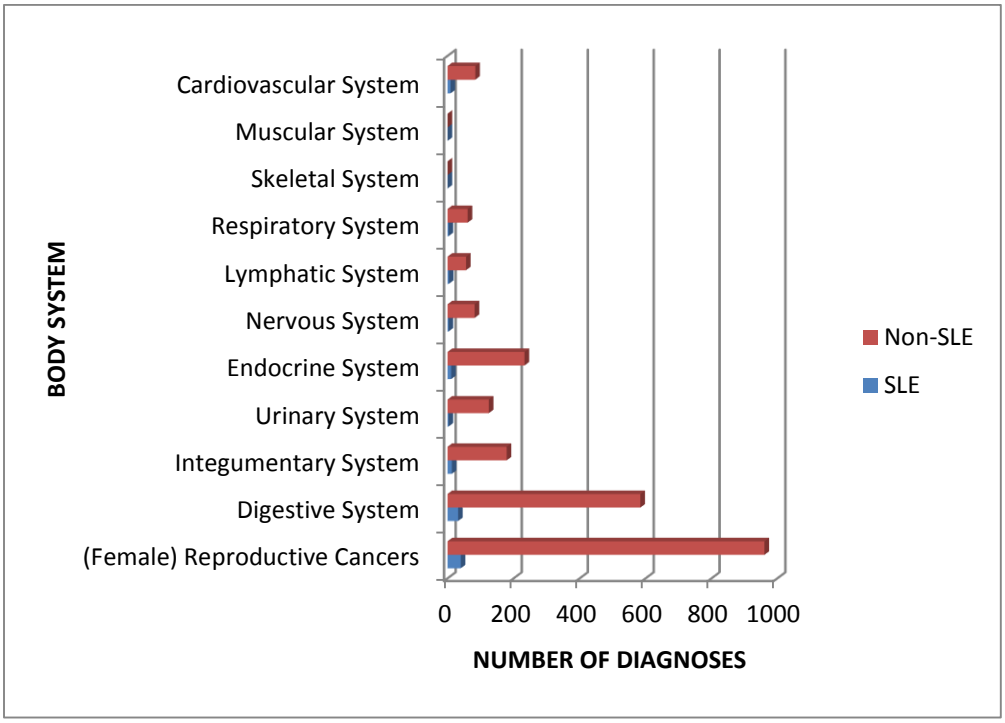

Figure 8. Number of malignancies by body system

Results

A binary logistic model was used to describe the fit for using the covariates to select the SLE and non-SLE cohorts found in CPRD. IPTW was the weight variable and there were two levels of responses, yes (has cancer) or no (does not have cancer). The regression parameters were estimated using Fischer's scoring method. The total number of observations read and used was 183,580. The unweighted SLE cohort yielded 3025 patients and 180,555 patients in the non-SLE cohort (Table 12).

Table 12

Unweighted SLE (1) and Non-SLE (0) Cohorts in CPRD

Table of LUPUS by CANCERFLAG				
		CANCERFLAG(CANCRFLAG)		
		0	1	Total
LUPUS(LUPUS)				
0	Frequency	178 98	2157	180555
1	Frequency	2906	119	3025
Total	Frequency	181304	2276	183580

In the unweighted cohorts, the relative risk of developing cancer due to having SLE was 3.39 (see Table 13).

Table 13

Statistics for Table of LUPUS by CANCERFLAG (unweighted)

Estimates of the Relative Risk (Row1/Row2)			
Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	3.3868	2.8059	4.088

Once the computer generated weights were applied, the SLE cohort weight was 18055.7 and 645580 in the non-SLE cohort (Table 14).

Table 14

IPTW weighted SLE (1) and Non-SLE (0) Cohorts in CPRD

		CANCERFLAG (CANCERFLAG)		
		0	1	Total
LUPUS(LUPUS)				
0	Frequency	195565	2472.89	198037
1	Frequency	450015	15582.9	465598
Total	Frequency	645580	18055.7	663636

In the weighted cohorts, the unadjusted relative risk of developing cancer due to having SLE is 2.74 (see Table 15).

Table 15

Statistics for Table of LUPUS by CANCERFLAG (weighted)

Estimates of the Relative Risk (Row1/Row2)			
Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	2.7385	2.6238	2.8581

An assessment of the fit of the binary logistic model against the data revealed that it was a good fit as demonstrated by the values of the intercept only versus the intercept and covariates in the three methods used to assess the model fit. Akaike Information Criterion (AIC) was used for the comparison of non-nested models on the same sample (University of California Los Angeles. 2015). The Schwarz Criterion (SC) and the AIC both penalize for the number of predictors in the model and the smallest SC and AIC are the most desired model (University of California at Los Angeles. 2015). The -2 Log L is used to test hypotheses in nested models and there is no real value in the numbers (UCLA. 2015). The intercept only column represents the response variable with no predictors in the model whereas the intercept and covariates column represents criterion statistics for the fitted model, which includes all independent models and the intercept (Table 16) (UCLA. 2015). For each of the criterion used, the lower value of the intercept and covariates versus the intercept only confirms that the binary logistic model used was a good fit for the data.

Table 16

Model Fit – SLE & Non-SLE (unweighted)

Criterion	Intercept Only	Intercept and Covariates
AIC	165773.2	163109.3
SC	165783.3	163129.6
-2 Log L	165771.21	163105.3

With no adjustment for age, the predictor variable of SLE indicates that a patient with SLE is 2.7 times more likely to develop cancer than is a non-SLE patient with 95% Confidence Limits (Table 17).

Table 17

Odds Ratio Estimates-SLE

Effect	Point Estimate	95%Wald Confidence Limits		Pr > ChiSq
LUPUS	2.738	2.623	2.857	<.0001

The fit of the binary logistic model was assessed with the variables of age. The binary logistic model was again a good fit as demonstrated by the lesser value of the intercept and covariates versus intercept only value for each of the three methods used to assess the model fit, AIC, SC, and -2 Log L (see Table 18). The greater difference in the intercept and the intercept wand covariates indicates that with the addition of age, the model was an even better fit than without the weighted covariates.

Table 18

Model Fit – SLE and Age

Criterion	Intercept Only	Intercept and Covariates
AIC	165773.2	149047.2
SC	165783.3	149077.5
-2 Log L	165771.21	149041.2

With adjustment for age, the predictor variable of SLE indicates that a patient with SLE is still 2.7 times more likely to develop cancer than is a non-SLE patient with 95% confidence limits (Table 19). Age was not a significant factor in whether the SLE patients developed cancer.

Table 19

Odds Ratio Estimates-SLE & AGE

Effect	Point Estimate	95% Wald Confidence Limits		Pr > ChiSq
LUPUS	2.669	2.556	2.788	<.0001
AGE	1.052	1.051	1.053	<.0001

All of the original covariates related to cancer development were not identifiable in CPRD because of the limitations of using a database that was designed to collect universal health information in the United Kingdom.

Odds ratio estimates were run to include age, contraception use, pregnancies, obesity, and smoking history (Table 20). Having had a pregnancy had the greatest effect on whether cancer developed when all the other variables were held constant. The next

greatest effect on cancer development was having SLE, followed by obesity, use of oral contraception, age, and then smoking. It was quite surprising that smoking actually seemed to have a protective effect on cancer development. These effects are stated with 95% confidence limits that with repeated trials; we would obtain the same results for each covariate. No additional post-hoc analyses were performed.

Table 20

Odds Ratio Estimates for Age, Contraception Use, Pregnancies, Obesity, and Smoking History

Effect	Point Estimate	95% Wald Confidence Limits		Pr > ChiSq
Lupus	2.906	2.781	3.035	<.0001
Age	1.06	1.058	1.061	<.0001
Contraception	1.602	1.521	1.687	<.0001
Pregnancy	3.146	2.98	3.322	<.0001
Obesity	2.476	2.4	2.555	<.0001
Smoking history	0.793	0.765	0.823	<.0001

Summary

This study sought to assess the association of cancer development in patients with SLE. In evaluating research question one, the risk of cancer development is increased in SLE patients compared to non-SLE patients. In the weighted cohorts, the unadjusted relative risk of developing cancer due to having SLE was 2.74. The greatest number of malignancies was female specific cancers such as breast, uterus and cervix followed by malignancies of the digestive system. This finding is congruent with findings from previous research, that several cancers are thought to be increased in people with SLE as

compared to persons without SLE (Bernatsky et al., 2005; Kiss et al., 2010; Parikh-Patel et al., 2008).

The balance created by using the IPTW to select the SLE and non-SLE cohorts attributed to a more accurate comparison of the risk of developing cancer between the two groups. The exposure to similar disorders and concomitant medications controlled for exposures that could have attributed to the risk of cancer development and biased the assessment of the cancer risk. Weighting the variables to control the number and similarity of variables in the patients in each cohort contributed to the balance of the variables between the study participants and results in less bias that could be attributed to medical history.

When assessing cervical cancer risk in SLE patients as compared to non-SLE populations, the null hypothesis failed to be rejected. The number of cervical cancers found in the non-SLE cohort (n = 88) were significantly greater than the number found in the SLE cohort (n = 2). This finding appear to contradict findings by Kiss et al. (2010) and Bernatsky et al. (2005), who reported cervical cancer to be increased in SLE patients.

The findings may appear to oppose previous findings because of the size of the CPRD database versus the sample size used by Kiss et al. (2010) and Bernatsky et al. (2005). This study had 3,025 SLE patients and 180, 555 non-SLE patients; Kiss et al. (2010) had only 860 patients with SLE and Bernatsky et al. (2005) had only 1,545 patients with SLE. Both studies compared the rates found in their SLE patients to rates found in the general population. In addition, in the Bernatsky et al. (2005) study, they noted that women with SLE have an increased risk of cervical dysplasia and atypia on

Pap testing as compared to non-SLE females. Cervical dysplasia and atypical pap tests do not always result in cervical cancer. This study compared only primary cervical cancer, not secondary malignancies, nor abnormal Pap tests. This study also compared non-SLE females in CPRD only and not in the general population. The patients were also balanced in terms of the types of other diagnoses and concomitant medications to which the two groups had been exposed. These factors all contribute to the seemingly contradictory findings in this study as compared to some previous studies.

In this chapter, deviations from the initial plan to use preselected covariates based upon the literature review were explained along with the rationale for the use of an alternative plan. Details of methods used to select and balance the two cohorts from CPRD were explained in detail. Checks to ascertain that the models used were included to confirm the appropriate fit of the model. Finally, findings from the various assessments that were performed to assess the association of cancer development in SLE patients as compared to non-SLE patients were presented. In Chapter 5, I will discuss the social change implications of these findings, the limitations of this study, and future recommendations for continued research to assess the association of cancer development in SLE as compared to non-SLE patients.

Chapter 5: Discussion, Conclusions, and Recommendations

Introduction

The purpose of this study was to assess whether there is an association between cancer developments in patients with SLE as compared to non-SLE patients. A propensity, score-matched, retrospective, cohort study among SLE and non-SLE female patients identified in CPRD was used. Female SLE patients were matched with non-SLE patients in CPRD. The matching was performed by using a high dimensional propensity model. The high dimensional propensity model was used to assign weights to variables based upon the analyses of common variables related to SLE and found in the database (Schneeweiss et al., 2009). The covariates used to select the SLE and non-SLE cohorts were diagnoses that were most commonly made in SLE patients and most commonly associated prescriptions made to patients with SLE. The selected covariates were chosen based on the frequency that they were found in the database for the selected population, in this case SLE. The covariates, concurrent diagnoses and concomitant medications, were then assimilated into a propensity score-based confounder adjustment model. The application of an IPTW to the SLE and non-SLE cohorts rendered a SLE cohort of 3,025 and a non-SLE cohort of 180,555 patients.

Previous research in this area has been based upon small study samples, cohorts that were not closed, which could increase the number of patients lost to follow-up, and the studies have lacked definitive diagnosis dates for SLE in the participants (Bernatsky et al., 2005; Parikh-Patel et al., 2008). As a result, previous studies were generally underpowered to be able to determine conclusively whether a relationship exists between

SLE and cancer (Bernatsky et al., 2005). In this study, I used a large, population-based database to test the relationship between SLE and cancer.

Use of the high dimensional propensity model allowed the cohorts in the study to be much better balanced than some of the studies that have been done in the past. The balance that was created by allowing the computer to search the entire CPRD database and select the most common diagnoses and concomitant medications that were assigned to patients with a diagnosis of SLE decreased the potential for bias that could result from selecting the non-SLE cohort based of age or age group. The balance in matching the cohorts decreased the potential for bias that could result due to the use of certain medications. The significant finding in the weighted cohorts was that the relative risk of developing cancer due to having SLE was 2.74 (95% CL). Another significant finding was that age was not a significant factor in whether the SLE patients developed cancer.

Interpretation of the Findings

The findings from this study confirm that cancers are increased in patients with SLE as compared to persons without SLE. In this study, the greatest number of malignancies was female-specific cancers such as breast, uterus, and cervix. The next largest group of malignancies was in the digestive system. The cohorts in this study were restricted to female only, and for that reason the fact that the greatest types of cancers were female-specific can neither confirm nor refute previous findings regarding the types of cancers increased in SLE patients.

As reported by the CDC (2014), age greater than 65 years is a risk factor for most types of cancers. As a person ages, he or she is exposed to the multitude of elements,

including both environmental and lifestyle factors, which are associated with increased risk of general cancer development (CDC, 2014). This study could not confirm or disconfirm whether age greater than 65 years resulted in a greater risk of developing cancer. I did find it increasing that age did not increase the development of cancer. This finding is in contradiction to findings by Extermann (2000) who found that increasing age increases the likelihood of developing cancer of some type because of the longer exposure to potential carcinogens in the environment, foods, and other factors.

Being overweight increases the risk of some cancers because of increased levels of estrogens and insulin (Cancer Research United Kingdom, 2014). OR estimates obesity had a 2.5 increased effect on whether cancer developed when all the other variables were held constant. A full-term pregnancy before the age of 17 years as compared to a woman who had her first pregnancy after the age of 25 years and women who have more than three pregnancies were associated with an increase in the risk of cervical cancer development (CDC, 2014). These two factors increase the risk of cervical cancer development because they increase the chances of acquiring a HPV infection because of the potential of the woman having unprotected sex with a greater number of sexual partners (ACS, 2014). The hormonal changes associated with pregnancy may also contribute to a weakened immune system and render the woman to susceptible to HPV infection (ACS, 2014). The OR estimates that if the patient had a pregnancy, she has a 3.1 (95% CL) increased chance of cancer development when all the other variables were held constant. In this study, any pregnancy was reported and not the number of

pregnancies. For this reason, the ability to assess the impact that pregnancy has on cancer development was not possible.

Use of estrogen-progestagen OCs over a 5-year or greater period may increase the development of cervical cancer (CDC, 2014). Although HPV exposure is known to be the most important cause of cervical cancer, when combined with long-term use of OCs, the RR of cervical cancer development was increased as duration of use increased. Smith et al. (2003) studied women with cervical cancer with no history of OC use (RR 1.9-2.2) compared to women who had used OCs (RR 1.6-3.9) for more than 10 years and found that the incidence of cervical cancer increased with longer use of OCs. OR estimated that contraception use had 1.6 increased effects on whether cancer developed, when all the other variables were held constant. In this study, the information regarding the amount of time that the patient used an OC and whether it was estrogen progestogen-based was not assessable; therefore, it is impossible to accept whether OC use affects the risk of cancer development.

Smoking renders a person exposed to many cancer-causing chemicals that affect multiple body organs when the chemicals are carried via the blood system to the organs. The chemicals act to damage the DNA of cells and may contribute to the development of multiple cancers (ACS, 2014). A history of tobacco use was a cancer risk factor in this study and was measured as either yes or no. OR estimated that smoking had a protective effect on whether cancer developed when all the other variables were held constant. This outcome was not expected because it is not consistent with the knowledge base. Neither CPRD nor HES provided enough details on smoking history to be able to assess this

variable appropriately. In this study, the amount of time that the patient smoked was not assessable; therefore, it is impossible to accept how smoking affects the risk of cancer development based upon the results from this study.

Limitations of the Study

The validity of this study depended on the accuracy and reliability of the information entered into the database. SLE is a difficult diagnosis to make and there may have been a risk of including non-SLE patients. The most common area of bias occurred in the collection of history of the patients admitted into the study. The study population included females with SLE and female non-SLE patients. Patients with a diagnosis of SLE were matched to non-SLE females using a high-dimensional propensity score. This method of using a high-dimensional propensity model is one method to address bias that can enter a study because of the history of the patients in the comparison groups. This method decreased bias that could present itself due to the selection process. A lack of consideration to the history of the patients, especially those factors that could lead to the development of SLE decreased bias between the comparison groups. The propensity score method of matching also addressed bias that was commonly a result of inclusion and exclusion criterion for a study. Selection using propensity score matching ensures that the comparison groups are equal in their covariates that could lead to the development of SLE and cancer.

The use of data that had already been collected has limitations in that the research questions must be tailored around the information that is available in the database. An additional limitation associated with using CPRD could be whether the data that were

entered into the database were entered with accuracy and reliability. There is no expectation that there is a difference in the accuracy of reporting between SLE and non-SLE patients. The Hospital Episode Statistics database was used for this study. The validity of this study depended on the accuracy and reliability of the information entered into the database. The most common areas of bias occur in the history of the patients admitted into the study, the inclusion criterion used to select patients into the study, and the methods used to analyze the study results (Gerhard, 2008).

Recommendations

The primary value of this study is that it used a large database to incorporate the numerous factors that are thought to lead to the development of SLE and cancer, both chronic diseases. There are many challenges in determining associations especially in chronic diseases such as SLE and cancer. These two diseases involve multiple factors that interact and result in their disease state. Consideration of the various factors that contribute to the two diseases is necessary to measure an association of the two diseases. The fact that both cancer and autoimmune diseases have been associated with diet, air quality, exposure to certain drugs, and personal habits makes the study of an association between the two diseases somewhat challenging. Studies completed to date have not shown consistent relationships of cancer development in patients with SLE.

In this study, careful consideration was given to the multiple factors that play a role in the development of SLE and in cancer. A large, population-based database was used to assess the relationship between SLE and cancer in a large population. The design of this study, and the methods to assess the findings were dynamic enough to allow for a

detailed analysis that was easily understood. A propensity score-matched retrospective cohort study among SLE and non-SLE female patients identified in CPRD was used to assess the association of cancer development in patients with SLE will be conducted. Non-SLE patients were matched with SLE patients in CPRD and then linked to an additional database for information on the covariates.

Balance between the SLE and non-SLE cohorts was obtained by using variables that were associated with patients with SLE to determine the treatment and non-treatment cohorts. This method of cohort selection balanced the SLE and non-SLE cohorts and balance between the study participants. I recommend that more studies to assess the association of cancer development in SLE patients be conducted utilizing a propensity score model to balance the study groups.

I further recommend that a database that contains more data that can be used to assess disease severity be used to study the association of cancer in patients with SLE. A focused medical history should be collected so that history such as age at first pregnancy and number of pregnancies, use of oral contraceptives, tobacco use, and viruses such as EBV/HPV can be assessed. Details of the medical history are essential in studying two chronic non-communicable diseases such as SLE and cancer.

Implications

Positive social change refers to involvement in activities that make improvements in the lives of individuals and communities locally and around the world. The goal of social change is to incorporate strategies that allow the individuals in the target population to maintain their dignity and self-worth. Positive social change results in

improvements in the health of the target population as well as their overall quality of life. The most important aspect of positive social change is that it gives individuals and eventually groups the power to improve the world around them (Walden University, 2011). My research, to assess the association of cancer development in persons with SLE disease, could promote positive social change by providing a better understanding of variables that may impact cancer development in patients with SLE. Understanding variables that can modify cancer risk can provide insight to factors that could possibly be altered to decrease the development of cancer in persons with SLE.

An understanding of these associations provides valuable insight to factors that can be altered to decrease the development of cancer in SLE patients. A matched retrospective cohort study among SLE and non-SLE patients was conducted using the propensity score methodology to help balance the differences between the comparison groups. The propensity score methodology created a similar distribution of observed baseline covariates between the two groups. The study outcomes could be used to promote positive social change by reinforcing current recommendations for cancer screenings in persons with SLE, which could enhance the ability to detect cancer early enough to decrease mortality due to cancer in persons with SLE.

Stronger adherence to cancer screening recommendations could enhance the ability to detect a cancer early enough so that treatment can be implemented that may result in a higher likelihood to effectively eradicate the cancer and decrease mortality due to cancer in persons with SLE. Results from this study could also equips persons with SLE with scientifically-based knowledge that may enable them to make decisions

regarding their care with a clearer understanding of the cancer risks inherent to persons with SLE, particularly when factored with their knowledge of their personal familial risks for cancer development.

Conclusion

In conclusion, better control of SLE flares has resulted in decreased deaths due to SLE activity and people affected by the disease are living longer. The increase in the lifespan of people with SLE has now shown that other chronic diseases are often the cause of death for persons with SLE. As new models of studying diseases are developed, they should be utilized as appropriate. Newer research models can assist in a better understanding of the multifactorial causations of chronic diseases since they often are a result of overlapping environmental, lifestyle, and everyday exposures. The economic burdens caused by autoimmune diseases and cancers make them a major public health concern. The increasing health-care burdens of both diseases must be better understood so that improved and targeted programs to ease the economic burdens on the public health system can be developed and implemented.

References

Abu-Shakra, M., Ehrenfeld, M., & Shoenfeld, Y. (2002). Systemic lupus erythematosus and cancer: Associated or not? *Lupus, 11,* 137-144. doi:10.1191/0961203302lu182rr

Achenza, M. I. S., & Selmi, C. (2012). Autoimmunity and cancer. *Asian Pacific Journal of Cancer Prevention, 13,* 29-40. Retrieved from http://dx.doi.org/10.7314/APJCP.2012.13.KKSuppl

American Cancer Society. (2014). *Lymphoma.* Retrieved from http://www.cancer.org/cancer/lymphoma/index

American College of Rheumatology. (2013). *The lupus initiative*. Retrieved from http://www.rheumatology.org/

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46,* 399-424. doi:10.1080/00273171.2011.568786

Azab, N. A., Bassyouni, I. H., Emad, Y., Abd Ei-Wahab, G. A., Hamdy, G., & Mashahit, M. A. (2008). CD4+CD25+ regulatory T cells (TREG) in systemic lupus erythematosus (SLE) patients: The possible influence of treatment with corticosteroids. *Clinical Immunology, 127,* 151-157. doi:10.1016/j.clim.2007.12.010

Bei, R., Masuelli, L., Palumbo, C., Modesti, M., & Modesti, A. (2009). A common repertoire of autoantibodies is shared by cancer and autoimmune disease patients: Inflammation in their induction and impact on tumor growth. *Cancer Letters, 281,* 8-23. doi:10.1016/j.canlet.2008.11.009

Belkaid, Y., Piccirillo, C. A., Mendez, S., Shevach, E. M., & Sacks, D. L. (2002). CD4+CD25+ regulatory T cells control Leishmania major persistence and immunity, *Nature, 420,* 502-507. doi: 10.1038/nature01152

Bengtsson, A. A., Rylander, L., Hagmar, L., Nived, O., & Sturfelt, G. (2002). Risk factors for developing systemic lupus erythematosus: A case-control study in southern Sweden. *Oxford Rheumatology, 41,* 563-571. doi.org/10.1191/096120301678416079

Bernatsky, S., Boivin, J. F., Joseph, L., Rajan, R., Zoma, A., Manzi, S., . . . Clarke, A. (2005). An international cohort study of cancer in systemic lupus erythematosus. *Arthritis & Rheumatism, 52*(5), 1481–1490. doi:10.1002/art.21029

Bernatsky, S., Ramsey-Goldman, R., & Clarke, A. (2006). Malignancy and autoimmunity. *Current Opinion Rheumatology, 18,* 129-134. doi.org/10.1097/01.bor.0000209423.39033.94

Bernatsky, S., Joseph, L., Boivin, J. F., Gordon, C., Urowitz, M., Gladman, D., . . . Clarke, A. E. (2008). The relationship between cancer and medication exposures in systemic lupus erythematosus: A case-cohort study. *Annals of the Rheumatic Diseases, 67*(1), 74-79. doi.org/10.1136/ard.2006.069039

Bertsias, G., Cervera, R., & Boumpas, D. T. (2012). Systemic lupus erythematosus: Pathogenesis and clinical features. *The European League Against Rheumatism*, 476-505. doi: 10.1136/annrheumdis-2012-201940

Broadbent, A. (2009). Causation and models of disease in epidemiology. *Studies in History and Philosophy of Biological and Biomedical Sciences, 40,* 302-311. doi:10.1016/j.shpsc.2009.09.006

Caliendo, M., & Kopeinig, S. (2005). Some practical guidance for the implementation of propensity score matching. *Discussion Paper Series Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor, IZA DP No. 1588,* 1-32.

Cancer Research United Kingdom. (2014). *The link between cancer and infections.* Retrieved from http://scienceblog.cancerresearchuk.org/2014/02/26/the-link-between-cancer-and-infections

Centers for Disease Control and Prevention. (2014). *Systemic lupus erythematosus.* Retrieved from http://www.cdc.gov/arthritis/basics/lupus.htm

Chang, E. T., Smedby, K. E., Hjalgrim, H., Schöllkopf, C., Porwit-MacDonald, A., Sundström, C., . . . Glimelius, B. (2005). Medication use and risk of non-Hodgkin's lymphoma. *American Journal of Epidemiology, 162,* 965-974. doi:10.1093/aje/kwi311

Clinical Practice Research Database. (2013). *Welcome to the clinical practice research datalink*. Retrieved from http://www.cprd.com/home/

Cooper, G.S., Dooley, M. A., Treadwell, E. L., St. Clair, E. W., & Gilkeson, G. S., .(2002). Hormonal and reproductive risk factors for development of systemic lupus erythematosus: Results of a population-based, case-control study. *Arthritis Rheumatology, 46*, 1830-1839. doi:10.1002/art.10365

Cox, D. R. (1972). Regression models and life-tables, *Journal of the Royal Statistical Society. Series B (Methodological), 34*, 187-220. Retrieved from http://www.jstor.org

Creswell, J.W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Cristaldi, E., Malaguarnera, G., Rando, A., & Malaguarnera, M. (2011). Possible link between autoimmunity and cancer. In C. Mavragani (Ed.), *Autoimmune disorders-pathogenic aspects* (pp. 387-417). Rijeka, Croatia: InTech. doi:10.5772/23988

Curtis, L. H., Hammill, B. G., Eisenstein, E. L., Kramer, J. M., & Anstrom, K. J. (2007). Using inverse probability-weighted estimators in comparative effectiveness analyses with observational databases. *Medical Care, 45*(10), S103-S107. doi:10.1097/MLR.0b013e31806518ac

Engels, E. A., Cerhan, J. R., Linet, M. S., Cozen, W., Colt, J. S., Davis, S., . . . Hartge, P. (2005). Immune-related conditions and immune-modulating medications as risk factors for non-Hodgkin's lymphoma: A case-control study. *American Journal of Epidemiology, 162*, 1153-1161. doi:10.1093/aje/kwi341

Extermann, M. (2000). Measuring comorbidity in older cancer patients. *European Journal of Cancer, 36,* 453-471. Retrieved from http://dx.doi.org/10.1016/S0959-8049(99)00319-6

Fairweather, D., & Rose, N. R. (2004). Women and autoimmune diseases. *Emerging Infectious Diseases, 10,* 2005-2011. doi:10.3201/eid1011.040367

Frankfort-Nachmias, C., & Nachmias, D. (2008). *Research methods in the social sciences* (7th ed.). New York, NY: Worth.

Gerhard, T. (2008). Bias: Considerations for research practice. *American Society of Health-System Pharmacists, 65,* 2159-2168. doi:10.2146/ajhp070369

Gill, J. M., Quisel, A. M., Rocca, P. V., & Walters, D. T. (2003). Diagnosis of systemic lupus erythematosus. *American Family Physician, 68,* 2179-2187. Retrieved from www.aafp.org/afp

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychology Methods, 15*, 234-249. doi:10.1037/a0019623.

Health & Social Care Information Center. (2014). *Hospital episode data.* Retrieved from http://www.hscic.gov.uk/hes

Heinze, G., & Jüni, P. (2011). An overview of the objectives of and the approaches to propensity score analyses. *European Heart Journal, 10,* 1-5. doi:10.1093/eurheartj/ehr031

Hildalgo-Conde, A., Liger, M.de H., Abarca-Costalago, M., Pérez, M. A., Valdivielso-Felices, P., González-Santos, P., & Fernández-Nebro, A. (2013). Incidence of cancer in a cohort of Spanish patients with systemic lupus erythematosus. *Reumatología Clínica, 9*, 359-364. doi:10.1016/j.reuma.2012.10.015

Kiss, E., Kovacs, L., & Szodoray, P. (2010). Malignancies in systemic lupus erythematosus. *Autoimmunity Reviews, 9*, 195-199. doi:10.1016/j.autrev.2009.07.004

Le, C. T. (2009). *Health and numbers: A problems based introduction to biostatistics* (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Lee, W. S. (2013). Propensity score matching and variations on the balancing test. *Empirical Economics*, *44*(1), 47-80. doi:10.1007/s00181-011-0481-0

Lewis, J. D., Bilker, W. B., Weinstein, R. B., & Stron, B. L. (2005). The relationship between time since registration and measured incidence rates in the General Practice Research Database. *Pharmacoepidemiology and drug safety, 14*, 443-451. doi:10.1002/pds.1115

Liang, J., Sun, L., Yeh, J., Lin, W., Chang, S., Sung, H., & Kao, C. (2012). Malignancies associated with systemic lupus erythematosus in Taiwan: A nationwide population-based cohort study. *Rheumatology International*, *32*, 773-778. doi:10.1007/s00296-010-1684-y

Manzi, S. (2009). Lupus update: Perspective and clinical pearls. *Cleveland Clinic Journal of Medicine, 76*, 137-142. doi:10.3949/ccjm.76a.gr005

Nagy, G., Koncz, A., & Perl, A. (2005). T and B cell abnormalities in systemic lupus erythematosus. *Critical Review in Immunology, 25*, 123-140. doi:10.1615/CritRevImmunol.v25.i2.30

Nakazawa, D. J. (2008). *The autoimmune epidemic: Bodies gone haywire in a world out of balance.* New York, NY: Touchstone/ Simon & Schuster.

National Cancer Institute. (2014). *Causes and prevention.* Retrieved from http://www.cancer.gov/

Najman, J. M. (1980). Theories of disease causation and the concept of a general susceptibility: A review. *Social Science & Medicine Medical Psychology Medical Sociology, 14*, 231-237. doi:10.1016/S0271-7123(80)91733-2

Nived, O., Bengtsson, A., Jönsen, A., Sturfelt, G., & Olsson, H. (2001). Malignancies during follow-up in an epidemiologically defined systemic lupus erythematosus inception cohort in southern Sweden. *Lupus, 10*, 500-504. doi:10.1191/096120301678416079

Parikh-Patel, A., White, R. H., Allen, M., & Cress, R. (2008). Cancer risk in a cohort of patients with systemic lupus erythematosus (SLE) in California. *Cancer Causes Control, 19*, 887-894. doi:10.1007/a10552-008-9151-8

Ragnarsson, O., Grondal, G., & Steinsson, K. (2003). Risk of malignancy in an unselected cohort of Icelandic patients with systemic lupus erythematosus. *Lupus, 12*, 687-691. doi:10.1191/0961203303lu443oa

Rassen, J. A., Glynn, R. J., Rothman, K. J., Setoguchi, S., & Schneeweiss, S. (2012). Applying propensity scores estimated in a full cohort to adjust for confounding in subgroup analyses. *Pharmacoepidemiology and Drug Safety, 21*, 697-709. doi:10.1002/pds.2256.

Rosenquist, R. (2008). Introduction: The role of inflammation, autoimmune disease and infectious agents in development of leukemia and lymphoma. *Journal of Internal Medicine, 264*, 512-513. doi:10.1111/j.1365-2796.2008.02028.x

Rudestam, K. E., & Newton, R. R. (2007). *Surviving your dissertation: A comprehensive guide to content and process* (3rd ed.). Thousand Oaks, CA: Sage.

SAEftyworks. (2015). *Database analytics automation software puts real-world evidence at your fingertips.* Retrieved from http://www.ubc.com/blog/database-analytics-automation-software-puts-real-world-evidence-your-fingertips

Schneeweiss, S., Rassen, J. A., Glynn, R. J., Avorn, J., Mogun, H., & Brookhart, M. A. (2009). High dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology, 20*, 512-522. doi:10.1097/EDE.0B013e318a663cc.

Schultz, D. R., & Harrington, W. J, Jr. (2003). Apoptosis programmed cell death at a molecular level. *Seminars in Arthritis & Rheumatism, 32*, 345-369. doi:10.1053/sarh.2003.50005.

Siegel, R., Ward, E., Brawley, D., & Jemal, A. (2011). Cancer statistics 2011: The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA: A Cancer Journal for Clinicians, 61*, 212-236. doi:10.3322/caac.20121.

Smith, J. S., Green, J., Gonzalez, B., Appleby, P., Peto, J., Plummer, M., . . ., Beral, V. (2003). Cervical cancer and use of hormonal contraceptives: A systemic review. *Lancet, 361*(9364), 1159-1167. doi:10.1016/S0140-6736(03)12949-2

Song, J. W., & Chung, K. C. (2010). Observational studies: Cohort and case-control studies. *Plastic Reconstruction Surgery. 126*, 2234-2242. doi:10.1097/PRS.0b013e3181f44abc

Statistics Solutions. (2014). *Statistical analysis system.* Retrieved from http://www.statisticssolutions.com/statistical-analysis-software-sas/

Statsdirect. (2014). *Conditional logistic regression.* Retrieved from http://www.statsdirect.com/help/default.htm#regression_and_correlation/conditional_logistic.htm

Sugihara, M. (2010). Survival analysis using inverse probability of treatment weighted methods based on the generalized propensity score. *Pharmaceutical Statistics, 9,* 21-34. doi:10.1002/pst.365

Tincani, A., Taraborelli, M., & Cattaneo, R. (2010). Antiphospholipid antibodies and malignancies. *Autoimmunity Reviews, 9*, 200-202. doi:10.1016/j.autrev.2009.04.001

Turesson, C., & Matteson, E. L. (2013). Malignancy as a comorbidity in rheumatic diseases. *Rheumatology, 52*, 5-14. doi:10.1093/rheumatology/kes189

University of California Los Angeles. (2015). *AIC: Reasons for AIC.* Retrieved from http://www-ssc.igpp.ucla.edu/personnel/russell/ESS265/Ch9/autoreg/node15.html#SECTION00043000000000000000

Walden University. (2011). *Social change*. Retrieved from https://www.waldenu.edu/about/so

Word, Z. H., & Matasar, M. J. (2012). Advances in diagnosis and management of lymphomas. *Blood and Lymphatic Cancer: Targets and Therapy*, *2012*(2), 29-55. Retrieved from http://dx.doi.org/10.2147/BLCTT.S15554

World Health Organization. (2014). *Cancer fact sheet.* Retrieved from http://www.who.int/mediacentre/factsheets/fs297/en/