Risk Factors for Breast Cancer among American Indian Women

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Melanie A Nadeau

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Jean Forster, DeAnn Lazovich

December 2016

Acknowledgements

First and foremost, I would like to thank creator, Gitchi Manitou, for providing me with the strength, health and wisdom to complete my doctoral dissertation. Thank you Gitchi Manitou for putting people in my path that could assist and guide me through the many obstacles I've encountered on this journey. A special thanks to my family. Words can't express how grateful I am to my mother (rest in peace), my father, my husband, my daughter, and my sister. Thank you for all of the sacrifices that you've made on my behalf and thank you for supporting and encouraging me throughout this experience. I would like to express my special appreciation and thanks to my community mentor, Dr. Lyle Best, my advisor, Dr. Jean Forster, and my co-advisor, Dr. DeAnn Lazovich. You have all been amazing mentors to me. Thank you so much for encouraging my research and for allowing me to grow as a researcher. Your advice on both research as well as on my career have been invaluable. And many thanks to everyone in my circle who also assisted me throughout the dissertation process. Thank you Dr. Donald Warne for providing time from work, Dr. Paula Carter for providing access to a biostatistician, Dr. Marilyn Klug for serving as my biostatistician mentor and the medical records staff at Indian Health Service for assisting me with my data collection efforts. And a huge thanks to all of my committee members. Thank you Dr. Jean Forster, Dr. DeAnn Lazovich, Dr. Alvaro Alonso and Dr. Susan Everson-Rose for serving on my committee and letting my defense be an enjoyable moment. All of you have been there to support me in countless ways and I will forever be grateful.

Dedication

I would like to dedicate this work to my beloved daughter Memphis Rose. Thank you so much baby bear for always being my biggest supporter. Your birth… your life… has driven me to want to be the best I can be. And I know someday you too will have this moment. I love you!

Abstract

Objective: This study was designed to identify risk factors associated with breast cancer among American Indian women, as the first step for developing a risk prediction model similar to the Gail model established for non-Hispanic white women.

Methods: A case-control study design was undertaken. Cases and controls were selected from among women undergoing mammograms at the Quentin N. Burdick Medical Care Facility (Indian Health Service) in Belcourt ND. For each woman with breast cancer (n=141), two controls were selected when possible (n=278). All women completed a breast cancer risk questionnaire at the time of their mammogram. This questionnaire was the primary source of data, supplemented by electronic and medical chart files. The risk factors examined were those included in the Gail model, including woman's age, age at first live birth, age of menarche, the number of previous benign breast biopsies, and the total number of first-degree relatives with breast cancer. In addition, body mass index and parity were also collected. Odds ratios and 95% confidence intervals were calculated using logistic regression.

Results: I did not find an association for American Indian women in North Dakota between most of the risk factors commonly identified in other populations and breast cancer. The majority of the associations were weakly positive with confidence intervals including the null value. Of all the risk factors examined, nulliparity was the only one that consistently showed a positive significant association.

Conclusion: Disparities in breast cancer incidence, mortality and screening among Northern Plains American Indians emphasize the need to better understand the risk factors associated with breast cancer in this population. It is my hope that this study will contribute to the development of a National Cancer Institute Breast Cancer Risk Assessment Tool that reflects the risk of breast cancer among American Indian women. Based on the results of my study, the value of risk prediction models in American Indian communities is uncertain and clinicians should be cautious in using the current Gail Breast Cancer Risk Assessment Tool to inform their American Indian patients of their risk for breast cancer.

Table of Contents

List of Tables

CHAPTER 1

I. INTRODUCTION AND SPECIFIC AIMS

Today most women can talk to their physician and find out what their risk of breast cancer is and what factors increase their risk of breast cancer by using an online assessment tool. The tool is called the Breast Cancer Risk Assessment Tool[1], also known as the Gail model. Risk factors included in the Gail model include woman's age, age at first live birth, age of onset of menstruation, the number of previous benign breast biopsies, and the total number of first-degree relatives with breast cancer.[2]

Breast cancer risk calculators assessing a woman's risk of developing breast cancer have been created for many populations including Whites[3–9], Hispanics[10], Asians[3,4,8,11], and African Americans.[3,4,8,12] A tool for American Indians does not exist despite the fact that breast cancer is the second most common cancer among American Indian women in the United States[13]. If a breast cancer risk calculator that is specific to American Indians is going to be created, information specific to American Indians is needed. For example, risk factors associated with breast cancer would need to be identified and their magnitude calculated, estimates of the age-specific breast cancer incidence rates for those with and without the identified risk factors would need to be obtained, and adjustments for competing causes of death would need to be made. To date, a breast cancer risk calculator does not exist for American Indians because information about risk factors for breast cancer for this population is extremely limited. As a result, the Breast Cancer Risk Assessment Tool uses data for White women when assessing the risk of breast cancer for American Indian women.[1]

The long-term goal of this study is to create a breast cancer risk prediction model that can estimate an American Indian woman's risk of developing breast cancer. This study focused on gathering the information needed to determine the relative risk for risk factors associated with breast cancer among American Indian women at the Quentin N. Burdick Medical Care Facility (Indian Health Service) in Belcourt ND. I explored the risk factors included in the Gail model as well as body mass index (BMI) and parity. My decision to include these risk factors was based on information from previous Gail model studies[2–12] (see Appendix 1) and a study by Chlebowski et al (see Appendix 1) which showed that breast cancer risk factor proportions (those included in the Gail model, BMI and parity) differ between Whites and American Indians/Alaska Natives.[14]

Using a case-control study design, I (1) collected information about characteristics that might put American Indian women at increased risk for breast cancer and (2) calculated the magnitude of risk factors associated with breast cancer in this population. This project is the first case-control study to examine risk factors for breast cancer among American Indian women who have been referred for mammographic screening. Additional studies are needed to accomplish the long-term goal of creating an American Indian breast cancer risk prediction model.

CHAPTER 2

II. BACKGROUND

Very little is known about the breast cancer risk profile among American Indian women. In 2005, Chlebowski et al examined data from the Woman's Health Initiative for 156,570 postmenopausal women recruited from 40 health facilities across the United States, ages 50 to 70 years with Whites representing 82% (n=129037) and American Indians/Alaska Natives representing 0.4% (n=696)[14]. The proportion of women with breast cancer risk factors by race/ethnicity was reported (Appendix 1, Table 1). Results showed that the proportion of women with known risk factors including age, age at first live birth, age at onset of menstruation, the number of previous benign breast biopsies, the total number of first-degree relatives with breast cancer, BMI, and parity differs between Whites and American Indians/Alaska Natives. In that study, American Indian women had a higher percentage of first degree relatives with breast cancer, a smaller percentage of individuals whose age at first occurrence of menstruation was younger than age 14, a lower percentage who had never given birth and a higher percentage of individuals with a BMI ≥30 compared to White woman.[14] However the small number of American Indians/Alaska Natives included (n=696, 11 of whom developed invasive breast cancer), limited the authors' ability to make conclusions about the associations of these factors with breast cancer.[14]

If a researcher is able to make conclusions about the association of potential risk factors with breast cancer, this information can then be used to create a breast cancer risk calculator. A breast cancer risk calculator is a mathematical model that is commonly used by physicians to determine a woman's risk of breast cancer. Physicians also use

information from these models to inform their patients about what risk factors are increasing their risk of breast cancer. Some of the more commonly known calculators are Gail, and Rosner and Colditz.[15] Other prediction models have been constructed by combining risk factors from these commonly known models.[15] Most models have been studied in White populations and some models have been studied in mixed populations with White being the majority.[15] Only a few models have been developed using data from minority populations, Asians and African Americans.[15] These models did not perform as well as other Gail models.[11,15]

A. The Gail Model

The Gail model, used most often by physicians, is available online at http://www.cancer.gov/bcrisktool/ and is known as the National Cancer Institute's Breast Cancer Risk Assessment Tool.[1] It is an interactive tool which helps to estimate a woman's risk of developing breast cancer. The user enters the necessary information and then the program immediately calculates the woman's breast cancer risk over the next five years, compared to the risk to women of the same age and ethnicity who have no risk factors.

The selection of risk factors used in the Gail model was determined from a nested case-control study conducted among participants in the Breast Cancer Detection Demonstration Project (BCDDP), which included five annual screening examinations for breast cancer conducted at 29 centers in the United States.[2] The risk factors included in the original Gail model were identified from both *in situ* (non-invasive) and invasive breast cancer cases.[2] *In situ* comprised about 10.3% of the cases. It is unclear why *in situ* cases were included in the original Gail model. The model includes the woman's

age, age at first live birth, age of onset of menstruation, the number of previous benign breast biopsies, and the total number of first-degree relatives with breast cancer.[2] These risk factors were determined as major predictors of risk for breast cancer by Gail et al and were also consistently identified in other studies.[2]

The Gail model has been useful in predicting risk of breast cancer in Whites[3–9], Hispanics[10], Asians[3,4,8,11], and African Americans.[3,4,8,12] The factors included in the various models for these subgroups are the same as the original Gail model; however the weight assigned to individual risk factors and the calculations used to assess the interactions of multiple factors have been adjusted to estimates specific to a particular subgroup.[1] For example, the model for Hispanics has been partially modified to include age-specific Hispanic breast cancer incidence rates and the model for African Americans has been modified to include both age-specific breast cancer incidence rates and relative risk for factors associated with breast cancer specific to the African American population. Although not included in Breast Cancer Risk Assessment Tool models, researchers have also modified the Gail model to include additional risk factors (Appendix 1, Table 2). However, including these additional factors only slightly improved the models so the factors were not included in the final model. It is also important to note that the Gail model is not always an accurate breast cancer risk assessment tool when applied in other communities. For example, researchers tested the validity of the Gail model in the Czech female population and found the Gail model was not an accurate tool for this population.[5]

B. Why I chose the Gail Model

I chose the Gail model to help accomplish the long-term goal of my study, which was to create a breast cancer risk prediction model that can estimate an American Indian

woman's risk of developing breast cancer, because the Gail model gives data driven breast cancer estimates in an easy-to-use format based on the answers to as few as four questions (in the first version of the model). It is also useful for studying and raising awareness of breast cancer. For example, the Gail model has been used to inform specific clinical decisions (whether or not to use tamoxifen to prevent breast cancer), plan intervention trials, and counsel women about their risk of disease.[16]

I also chose the Gail model because it performs better than other models. In 2011, Anothaisintawee et al conducted a systematic review of model performances for breast cancer risk prediction models.[15] The results from this review helped to identify the most reliable model and indicate the strengths and weaknesses of each model.[15] The Gail model was identified as the most reliable. It has also been in existence the longest, modified the most (see Appendix 1, Table 2), and validated by other scholars.[15] The Gail model is a good predictor of breast cancer at the population level but not at the individual level. The model is well calibrated in the sense that the expected numbers of breast cancers predicted for subsets of women agree well with the observed numbers of cancers that actually develop. However, this model (or any risk model) cannot reliably discriminate women who will develop breast cancer from those who will not and because of this, it (and all models of this type) has been criticized.[15] Calibration is a model's ability to correctly predict the number of events that develop in the cohort and subgroups of the cohort. Having a well calibrated model can be useful for early detection or prevention of cancer because it gives both the provider and the patient a general perspective on risk and a formal weighing of risks and benefits; for example, the

information could be useful for considering whether or not a woman in her 40s should have a mammogram.

Having a highly discriminatory model can be beneficial if a need exists to allocate prevention resources under cost constraints or to implement a high-risk prevention strategy. For example, when implementing interventions with adverse side effects, the investigator may want to target only those at high risk. One way to potentially increase a model's discriminatory accuracy is to include more risk factors.

C. Determining potential risk factors to include in my study

Given information resulting from the Chlebowski, et al, study on breast cancer risk factors in American Indians/Alaska Natives and previous research conducted on the Gail model, I proposed to explore the risk factors included in the Gail model as well as BMI and parity. I was interested in looking at these risk factors because uncertainty existed as to whether risk factors for breast cancer among White women are the same as for American Indian women. Furthermore, if associations could be found similar to White women, the magnitude of the associations between the risk factor and breast cancer may differ. Given that the model was adjusted for the same risk factors in other populations, I hypothesized that the risk factors would be the same in the American Indian population; however the magnitude of the associations would likely differ.

D. What information is needed to create an American Indian Gail Model?

To create a breast cancer risk calculator that is specific to American Indian populations, information specific to American Indians is needed. For example, the investigator must calculate the relative risk for factors associated with breast cancer in

American Indian populations, estimate the age-specific breast cancer incidence rates for those with and without the identified risk factors in American Indian population, and adjust for competing causes of death in American Indian populations. My study focused on the first step, namely, collecting data to determine the associations for risk factors for breast cancer in one American Indian population.

E. American Indian data limitations

Although estimating breast cancer incidence rates was not the focus of my study, it was necessary to use published incidence rates to determine the feasibility of my study. However, limitations must be considered when using cancer registry data to determine cancer rates among American Indians. Health status assessments for American Indians are often incomplete because of racial misclassification and lack of data collection efforts.[17] Racial misclassification occurs when health providers misclassify their patients' race. For example, an American Indian woman may be classified as White by her health provider because of her name, residence or appearance. In addition, complete cancer data for American Indians does not exist because the cancer registries only collect data for certain geographical areas which limits the generalizability of the data.[18] Another limitation is that American Indians represent the smallest population of any nationally reported subgroup which complicates statistical analysis. As of 2010, 5.2 million people (1.7%) in the US reported being American Indian/Alaska Native.[19]

Geographical differences also exist. The cancer burden varies significantly among the 564 federally recognized tribes across the United States. American Indians across the United States differ in culture, environmental exposures, and healthcare and these differences combine to produce unusual patterns of cancer occurrence.[20] Efforts to

obtain more accurate information are currently underway and include linking data between national cancer registries and the Indian Health Service (IHS).[21]

Limitations will also differ based on which data source is used. For example, National Cancer Institute data has different limitations than Indian Health Service data. Surveillance, Epidemiology, and End Results Program (SEER), a program of the National Cancer Institute, is a source of information on cancer incidence and survival in the United States. The cancer rates, considering all cancer sites, for American Indians/Alaska Natives from several SEER registries are considerably higher compared with the overall SEER rates by 2-fold to 5-fold even though the cancer rates for American Indian/Alaska Native populations generally have been reported as lower than the total population for the United States.[22] This is likely due to misclassification which is more likely to occur with SEER data compared to Indian Health Service data. American Indians who can show proof that they are enrolled members of a federally recognized tribe are eligible to receive healthcare, without charge, at the Indian Health Service which is comprised of federal and tribal hospitals and clinics. Healthcare provided by the IHS is most common in Contract Health Service Delivery Area (CHSDA) counties which are generally defined as counties that contain or are adjacent to federally recognized tribal reservations and/or trust lands.[17] Research shows that racial misclassification occurs less frequently in CHSDA counties[17] so IHS data are more inclusive of American Indians than SEER data. Furthermore, since cancer registries only collect data for certain geographical areas, many American Indian tribes are not included in SEER data collection efforts. However only about half of American Indians in the

U.S. live in CHSDA counties, and thus may not be served by IHS clinics or included in IHS data.

F. American Indian breast cancer incidence rates

Given these limitations, it is best to collect incidence data from the geographical area of interest when possible. According to data based on Indian Health Service Contract Health Service Delivery Areas (CHSDA) counties, the age-adjusted incidence for breast cancer among American Indians/Alaska Natives females was 91.9 per 100,000 compared to 128.1 per 100,000 for White females from 2008-2012 in the United States.[23] Breast cancer incidence rates in the Northern Plains Indian Health Service area among American Indian women were 112.6 per 100,000 compared to 125.5 per 100,000 for White women from 1999-2009.[24] The Northern Plains includes the following states: Illinois, Indiana, Iowa, Michigan, Minnesota, Montana, Nebraska, North Dakota, South Dakota, Wisconsin and Wyoming.[24] The age-adjusted breast cancer incidence for American Indian females in North Dakota was 100.8 per 100,000 compared to 125.9 per 100,000 for White females from 2009-2013.[25] The age-adjusted breast cancer incidence for American Indian females in Rolette County, North Dakota was 139.2 per 100,000 compared to 165.6 per 100,000 for White females from 2009-2013.[26]

G. In summary

American Indians have lower breast cancer incidence rates nationally, in the Northern Plains regional area, in North Dakota and in Rolette County, North Dakota compared to Whites. When exploring breast cancer incidence among American Indian women in the United States, it is important to be mindful of the many limitations of working with American Indian cancer data which include racial misclassification, limited

data collection, small population and geographical differences. It is also important to recognize that these limitations will vary by data source.

As the first case-control study conducted among American Indian women to assess breast cancer risk factors, this study focuses on gathering the information needed to determine the relative risk for potential risk factors associated with breast cancer among American Indian women at the Quentin N. Burdick Medical Care Facility (Indian Health Service) in Belcourt ND. Using a case-control study design, I assess the following specific aims: (1) establish which characteristics put American Indian women at increased risk for breast cancer and (2) calculate the magnitude of risk factors associated with breast cancer in this population. I explore the risk factors included in the Gail model as well as body mass index (BMI) and parity. Additional studies are needed to accomplish the long-term goal, which is to create an American Indian breast cancer risk prediction model.

CHAPTER 3

III. RESEARCH METHODS

This study was approved by the Indian Health Service Great Plains Institutional Review Board (IRB) and University of Minnesota IRB. See Appendix 2 for resolution in support of my proposed study that was passed by the Turtle Mountain Band of Chippewa, where the clinic is located, a letter of accommodation from the Indian Health Service CEO in Belcourt, ND, and human subjects' protection considerations. I also completed a background check and HIPAA, Privacy Act, and Information System Security trainings.

A. Setting

The study was conducted at the Quentin N. Burdick Memorial Health Care Facility (Indian Health Service) in Belcourt, ND. I chose the Quentin N. Burdick Memorial Health Care Facility primarily because Belcourt is home to the Turtle Mountain Band of Chippewa, the largest population of American Indians in North Dakota.[27] I also chose to conduct research in Belcourt because I am an enrolled member of the Turtle Mountain Band of Chippewa. Having previously conducted research in my community, I have many personal and professional contacts that made conducting the research feasible.

B. Design - Case/Control Study

A case-control study design was undertaken to assess the association between potential risk factors and breast cancer among a population of American Indian women who obtained their health care from Quentin N. Burdick Memorial Health Care Facility.

C. Inclusion criteria

Cases were women with pathologically confirmed incident invasive breast cancer diagnosed between 1985 and 2015 who had a mammogram prior to the date of diagnosis. Case ethnicity was defined as American Indian by virtue of their eligibility for care within the Indian Health Service. Controls were women in IHS files without breast cancer who had had a mammogram during the same timeframe as the cases. All women in my study were required to have had a mammogram to be eligible because it was the main source of the risk factor information.

D. Case/Control Identification Process

An electronic database, the Resource and Patient Management System, at the Indian Health Service was queried using ICD-9 code 174 to identify patients with a breast cancer diagnosis (cases) from May 1990 (when it was established) through January 29th, 2016, which was the end of data collection. When a clinician sees a patient, they may enter a diagnosis dated the date of the visit even though the diagnosis may be a historical diagnosis. For example, a patient's diagnosis may be dated in 1990 when they first see a new provider; however they may have been historically diagnosed by another provider 5 years earlier. This is why my list went back to 1990, however, after completing my chart review I have cases that were historically diagnosed back to 1985. A total of 170 patients with a possible breast cancer diagnosis were identified. A pathological diagnosis was confirmed for 141 (82.94%) patients. The remaining 29 were eliminated for the following reasons: unable to find risk factor data (n=1), unable to identify controls (n=2), unable to retrieve chart from archives (n=4), no previous mammogram (n=9), and unable to confirm a pathological diagnosis for breast cancer (n=13).

For each case, two controls were selected when possible (n=278) from women who had had a mammogram at the facility and had not had a breast cancer diagnosis. After all the cases had been identified, I frequency matched on age by selecting controls that were in the same 5-year age category as the cases. The two controls selected also had to have mammograms that were the closest to the date of the diagnostic mammogram of the case.

The first mammogram in the Belcourt IHS facility recorded electronically occurred on September 16, 1993. Controls for cases who had mammograms after September 16, 1993 (cases=120) were identified electronically. Controls for cases who had a mammogram prior to this date (cases=21) were identified using the following process:

- A list of all female patients seen at the Indian Health Service was printed in 5-year age groups. For example, women who were born from 1915 through 1919 were printed out on one list. Women born from 1920 through 1924 were on another list. Lists were printed for 5-year age groups through 1978, the year that the youngest case was born.
- Every woman on each list was given a number and these numbers were put in random order for each list.
- Starting with the first random number, I attempted to locate a radiology file in both the active and inactive section.
- If the file could not be found, I moved on to the next random number. This process continued until a control with a mammogram prior to September 16th, 1993 could be identified.

- Once identified, data was collected from the radiology, electronic and chart files.

E. Data Collection and Documentation

Since 1986, the Quentin N. Burdick Healthcare Facility asked every woman undergoing a mammogram to complete a brief questionnaire that included the following questions about risk for breast cancer: date of birth, age at first pregnancy, menstrual history - age at onset, ever had breast surgery (if yes, the patient is asked to specify mastectomy, biopsy, aspiration, other, right and/or left, and date), and if any blood relative had breast cancer (if yes, the patient is asked to specify mother, sister, grandmother, aunt, age of each relative identified and whether the relative was maternal or paternal). See Appendix 3 for an example of the questionnaire that is administered at this facility. These questionnaires, kept in the Radiology department, served as the primary source of data for the study.

Using the list of identified patients, I first attempted to locate the radiology file for each potential participant. Radiology files for patients who had not died were filed by chart number and easy to retrieve. Once deceased, a patient's radiology file was inactive, refiled by year of death and destroyed after 7 years. Mammograms forms were available for 110 (78%) cases and 278 (100%) controls. Each participant's mammogram form was reviewed. If the mammogram form was no longer available, the radiology chart was reviewed.

If the mammogram leading up to the date of diagnosis could be identified, the patient was included in the study as a case. Each study participant was assigned a unique study ID which was linked to their chart number on a list that was kept in a secure area at the Indian Health Service facility. Data was recorded for risk factors included in the

study. Date of previous benign breast biopsy, date of mastectomy and if the patient had an active radiology, electronic, and hard chart on file was also recorded (see appendix 4, Table 3 for an example of data collection sheet). I also noted whether review was completed for the radiology, electronic and medical chart files (see Table 1). Personal computers are not allowed in secured areas at the facility so all data was entered into an Excel spreadsheet off site. No identifying data was collected. Information on the mammogram questionnaire did not include BMI, age at first live birth and number of live births. When possible, BMI (the first BMI documented before and after mammogram) and number of live births was collected from the electronic medical record. Otherwise this information along with age at first live birth was collected from the chart.

F. Archived Charts

Indian Health Service retains medical records for 7 years after the last episode of care. Records are then retired to the Federal Records Centers (FRC) in Denver CO. A total of 39 charts were archived, 38 cases and 1 control, in the FRC. Charts archived at the FRC can be requested by the Indian Health Service medical records team. Medical records were requested five at a time and took approximately one month to receive. Once received, the records were reviewed and re-boxed for return. Medical records staff would then request the next batch of five.

Using this set of procedures, the study obtained data for 141 cases and 248 controls.

G. Missing data

After data collection was complete, the data were cleaned and missing data were imputed. Missing data occurred for both cases and controls. The same imputation

method was used to address missing data for both cases and controls. Percent missing for the variables was as follows: age at screening (n=0), age at first pregnancy (n=10, 2.36%), age at first live birth (n=63, 15.04%), age at onset of menstruation (n=8, 1.90%), number of previous benign breast biopsies (n=1, 0.24%), total number of first degree relatives with breast cancer (n=0), number of pregnancies (n=2, 0.47%), number of live births (n=25, 5.97%) and BMI (n=38, 8.96%). When age at first live birth was missing, age at first pregnancy was substituted. When age at onset of menstruation was missing, the average difference between age at onset of menstruation and age at first live birth, which was 7, was subtracted from age at first live birth to calculate age at onset of menstruation. When number of previous benign breast biopsies was missing, I looked at the average age for number of previous benign breast biopsies in my dataset. When the number of previous benign breast biopsies equaled 0, the average age was 54. When the number of previous benign breast biopsies equaled 1, the average age was 59. When the number of previous benign breast biopsies equaled 2, the average age was 63. When the number of previous benign breast biopsies equaled 3, the average age was 68. The one individual that was missing the number of previous benign breast biopsies was age 52 so I imputed 0 for their number of previous benign breast biopsies. When number of live births was missing, number of pregnancies was substituted. When BMI was missing, the following regression was used to fill in the missing: BMI=.763 casesness (whether or not the patient had breast cancer) + .104 number of live births + .027 age at onset of menstruation.

H. Categorization of variables and statistical analysis

Two approaches were taken to examine risk factors for breast cancer. In the first approach, risk factors and the categorization used in the Gail model were included. In the second approach, risk factors were categorized according to their distribution in the population of American Indian women.

The following Gail model categorization was used for Phase I of the analysis: age at screening (<50, ≥50), age at first live birth (<20, 20 to 24, 25 to 29, ≥30, never given birth), age of onset of menstruation (<12, 12 to 13, ≥ 14 years), the number of previous benign breast biopsies (0, 1, ≥ 2), the number of first-degree relatives with breast cancer (0, 1, ≥ 2), BMI (<25, 25 to 30, >30), number of live births (1, 2, 3, 4, ≥5, never given birth).

I also assessed the distribution of the data and proposed new categories for Phase II of the analysis. Compared to non-Hispanic white women, American Indian women in this study were screened at an older age and were younger at age of first live birth. They also had a lower number of previous benign breast biopsies and first-degree relatives with breast cancer and a higher BMI and number of live births compared to White women. After reviewing the categories with a physician, biostatistician and epidemiologist, I finalized the new categories for the study population and re-analyzed the data. The new categories used for Phase II of the analysis were as follows: age at screening (<56, ≥56), age at first live birth (<18, 18 to 19, 20 to 21, ≥22, never given birth), age of onset of menstruation (<12, 12, 13, ≥14 years), the number of previous benign breast biopsies (0, ≥ 1), the number of first-degree relatives with breast cancer (0, ≥ 1), BMI (<25, 25 to

29.99, 30 to 32.49, 32.50 to 34.99, ≥35), number of live births (1 to 2, 3 to 4, ≥5, never given birth).

I ran the following statistical models for both Phase I (original categories) and Phase II (new categories) of the analyses: (1) unconditional logistic regression for the crude association of age at screening, age at first live birth, age at onset of menstruation, the number of previous benign breast biopsies, the number of first-degree relatives with breast cancer, BMI, number of live births (all variables in the study), (2) multivariate unconditional logistic regression model – included woman's age, age at first live birth, age of onset of menstruation, the number of previous benign breast biopsies, and the total number of first-degree relatives with breast cancer (variables included in the Gail model), (3) multivariate unconditional logistic regression model – included woman's age, age at first live birth, age of onset of menstruation, the number of previous benign breast biopsies, the total number of first-degree relatives with breast cancer, and BMI (variables included in the Gail model plus BMI), (4) multivariate unconditional logistic regression model – included age at screening, age at onset of menstruation, the number of previous benign breast biopsies, the number of first-degree relatives with breast cancer, BMI, number of live births (all study variables except age at first live birth). The measure of association utilized was the odds ratio (OR). The OR is used to compare the relative odds of the occurrence of breast cancer given the potential risk factors. The 95% confidence interval was used to estimate the precision of the OR.

Table 1. Radiology, electronic and medical chart review - percent complete for cases/controls

	Cases	Controls
	n %	n %
Total n=419	141 (100.0)	278 (100.0)
Radiology review	110 (78.0)	278 (100.0)
Electronic		
review	102 (72.3)	277 (99.6)
Medical		
chart review	70 (49.6)	141 (50.7)
No Radiology review	31 (22.0)	0 (0.0)

CHAPTER 4

IV. RESULTS

In a univariate analysis of risk factors for breast cancer among American Indian women of the Northern Plains (Table 2), associations did not follow expected patterns as observed for White women. Most associations were positive, but often weakly so, and confidence intervals included the null value for the most part. Rather than an increased risk associated with older age at first live birth, odds ratios for each increasing age category relative to women in the youngest age group were inversely related to breast cancer. Only nulliparity was associated with a greater than two-fold increase risk of breast cancer, but confidence intervals were wide and included 1.0. Compared to women whose age of menstruation onset was 14 years or older, no trend of increased risk of breast cancer was observed with decreasing age at menstruation. Having 2 or more previous benign breast biopsies or two or more first degree relatives with breast cancer were associated with only weak to moderate likelihood of breast cancer, but the results were not statistically significant. Except for women who were nulliparous relative to women with 5 or more children (OR 2.59, 95% CI 1.08-6.20), no association nor trend was observed for parity or BMI and risk of breast cancer.

When a multivariate analysis restricted to the four risk factors included in the Gail model was conducted (Table 3), the association for breast cancer among women who were nulliparous in comparison to women who first gave birth before age 20 was strengthened and the confidence interval excluded the null value. Although the odds ratio for women whose age at onset of menstruation was 12 or 13 relative to women whose onset was age 14 was also stronger and now statistically significant compared to

the univariate analysis, a lack of a trend persisted. Inclusion of BMI in a multivariate model of Gail model risk factors did not alter these conclusions (Table 4). And when associations for breast cancer with age at onset of menstruation, benign breast biopsies, family history of breast cancer were simultaneously examined with BMI and parity, the statistically significant association for menstrual onset of 12 or 13 years remained, while the association for breast cancer among nulliparous women relative to women with 5 or more births strengthened and remained statistically significant (OR 3.02, 95% CI 1.23-7.40, Table 5), as was observed in the univariate analysis (Table 2). The c-statistic for all models in the Phase I analysis ranged from .59-.61.

For Phase II of the analysis, I proposed new categories. The new categories were as follows: age at screening (<56, ≥56), age at first live birth (<18, 18 to 19, 20 to 21, ≥22, never given birth), age of onset of menstruation (<12, 12, 13, ≥14 years), the number of previous benign breast biopsies (0, ≥ 1), the number of first-degree relatives with breast cancer (0, ≥ 1), BMI (<25, 25 to 29.99, 30 to 32.49, 32.50 to 34.99, ≥35), number of live births (1 to 2, 3 to 4, ≥5, never given birth). Phase II analysis and findings (Table 6-9) did not add anything new to Phase I findings. I did not find an association for American Indian women in North Dakota between most of the risk factors commonly identified in other populations and breast cancer. The majority of the associations were weakly positive with confidence intervals including the null value. Of all the risk factors examined, nulliparity was the only one that consistently showed a positive significant association. The c-statistic for all models in the Phase II analysis ranged from .61-.63.

Table 2. Phase I analysis (original categories) - unconditional logistic regression - assessing the association between each risk factor and breast cancer.

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<50	43	(30.5)	91	(32.73)		
≥50	98	(69.5)	187	(67.27)	1.11 (.72-1.72)	.64
Age at first live birth						
<20						
20 to 24	72	(51.06)	139	(50.00)		
25 to 29	48	(34.04)	106	(38.13)	.87 (.56-1.36)	.55
≥30	5	(3.55)	14	(5.04)	.69 (.24-1.99)	.49
Nulliparous	3	(2.13)	8	(2.88)	.72 (.19-2.81)	.64
	13	(9.22)	11	(3.96)	2.28 (.97-5.35)	.06
Age at onset of menstruation						
≥14	39	(27.66)	98	(35.25)		
12 to 13	79	(56.03)	129	(46.40)	1.54 (.97-2.45)	.07
<12	23	(16.31)	51	(18.35)	1.13 (.61-2.10)	.69
Number of previous benign breast biopsies						
0	111	(78.72)	221	(79.50)		
1	25	(17.73)	51	(18.35)	.98 (.57-1.66)	.93
≥2	5	(3.55)	6	(2.16)	1.66 (.50-5.56)	.41
Total number of first degree relatives with breast cancer						
0						
1	104	(73.76)	216	(77.70)		
≥2	31	(21.99)	53	(19.06)	1.22 (.74-2.01)	.45
	6	(4.26)	9	(3.24)	1.39 (.48-3.99)	.55
BMI, kg/m^2						
<25	15	(10.64)	29	(10.43)		
25-30	50	(35.46)	92	(33.09)	1.05 (.52-2.14)	.89
>30	76	(53.90)	157	(56.47)	.94 (.47-1.85)	.85
Parity (# of live births)						
≥5	47	(33.33)	103	(37.05)		
4	18	(12.77)	44	(15,83)	.90 (.47-1.71)	.74
3	28	(19.86)	64	(23.02)	.96 (.55-1.68)	.88
2	26	(18.44)	41	(14.75)	1.39 (.76-2.53)	.28
1	9	(6.38)	15	(5.40)	1.32 (.54-3.22)	.55
Nulliparous	13	(9.22)	11	(3.96)	2.59 (1.08-6.20)	.03

Table 3. Model I - Phase I analysis (original categories) - multivariate unconditional logistic regression model – includes risk factors included in the Gail model

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<50	43	(30.5)	91	(32.73)		
≥50	98	(69.5)	187	(67.27)	1.09 (.69-1.72)	.71
Age at first live birth						
<20						
20 to 24	72	(51.06)	139	(50.00)		
25 to 29	48	(34.04)	106	(38.13)	.84 (.53-1.31)	.44
≥30	5	(3.55)	14	(5.04)	.72 (.25-2.12)	.55
Nulliparous	3	(2.13)	8	(2.88)	.68 (.17-2.69)	.58
	13	(9.22)	11	(3.96)	2.54 (1.07-6.02)	.03
Age at onset of menstruation						
≥14	39	(27.66)	98	(35.25)		
12 to 13	79	(56.03)	129	(46.40)	1.62 (1.01-2.61)	.05
<12	23	(16.31)	51	(18.35)	1.14 (.61-2.15)	.68
Number of previous benign breast biopsies						
0	111	(78.72)	221	(79.50)		
1	25	(17.73)	51	(18.35)	.99 (.57-1.71)	.97
≥2	5	(3.55)	6	(2.16)	1.71 (.50-5.91)	.39
Total number of first degree relatives with breast cancer						
0						
1	104	(73.76)	216	(77.70)		
≥2	31	(21.99)	53	(19.06)	1.26 (.75-2.10)	.38
	6	(4.26)	9	(3.24)	1.53 (.51-4.51)	.44

Table 4. Model 2 - Phase I analysis (original categories) - multivariate unconditional logistic regression model – includes risk factors included in the Gail model plus BMI

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<50	43 (30.5)		91 (32.73)			
≥50	98 (69.5)		187 (67.27)		1.10 (.69-1.74)	.69
Age at first live birth						
<20						
20 to 24	72 (51.06)		139 (50.00)			
25 to 29	48 (34.04)		106 (38.13)		.83 (.53-1.31)	.43
≥30	5 (3.55)		14 (5.04)		.72 (.25-2.10)	.54
Nulliparous	3 (2.13)		8 (2.88)		.67 (.17-2.64)	.57
	13 (9.22)		11 (3.96)		2.52 (1.06-6.00)	.04
Age at onset of menstruation						
≥14	39 (27.66)		98 (35.25)			
12 to 13	79 (56.03)		129 (46.40)		1.67 (1.03-2.70)	.04
<12	23 (16.31)		51 (18.35)		1.18 (.62-2.24)	.61
Number of previous benign breast biopsies						
0	111 (78.72)		221 (79.50)			
1	25 (17.73)		51 (18.35)		.99 (.57-1.71)	.96
≥2	5 (3.55)		6 (2.16)		1.68 (.49-5.79)	.41
Total number of first degree relatives with breast cancer						
0						
1	104 (73.76)		216 (77.70)			
≥2	31(21.99)		53 (19.06)		1.26 (.75-2.10)	.38
	6 (4.26)		9 (3.24)		1.54 (.52-4.57)	.43
BMI, kg/m^2						
<25	15 (10.64)		29 (10.43)			
25 to 30	50 (35.46)		92 (33.09)		1.06 (.51-2.19)	.89
>30	76 (53.90)		157 (56.47)		.88 (.44-1.78)	.73

Table 5. Model 3 - Phase I analysis (original categories) - multivariate unconditional logistic regression model (includes all study variables except age at first live birth)

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<50	43	(30.5)	91	(32.73)		
≥50	98	(69.5)	187	(67.27)	1.14 (.70-1.85)	.60
Age at onset of menstruation						
≥14	39	(27.66)	98	(35.25)		
12 to 13	79	(56.03)	129	(46.40)	1.66 (1.03-2.70)	.04
<12	23	(16.31)	51	(18.35)	1.21 (.64-2.29)	.56
Number of previous benign breast biopsies						
0	111	(78.72)	221	(79.50)		
1	25	(17.73)	51	(18.35)	.99 (.57-1.72)	.98
≥2	5	(3.55)	6	(2.16)	1.66 (.49-5.73)	.43
Total number of first degree relatives with breast cancer						
0						
1	104	(73.76)	216	(77.70)		
≥2	31	(21.99)	53	(19.06)	1.25 (.75-2.09)	.40
	6	(4.26)	9	(3.24)	1.38 (.47-4.07)	.56
BMI, kg/m^2						
<25	15	(10.64)	29	(10.43)		
25-30	50	(35.46)	92	(33.09)	1.04 (.50-2.16)	.92
>30	76	(53.90)	157	(56.47)	.87 (.43-1.75)	.69
Parity (# of live births)						
≥5	47	(33.33)	103	(37.05)		
4	18	(12.77)	44	(15,83)	.97 (.49-1.89)	.92
3	28	(19.86)	64	(23.02)	1.04 (.58-1.87)	.90
2	26	(18.44)	41	(14.75)	1.46 (.78-2.72)	.24
1	9	(6.38)	15	(5.40)	1.32 (.53-3.29)	.55
Nulliparous	13	(9.22)	11	(3.96)	3.02 (1.23-7.40)	.02

Table 6. Phase II analysis (new categories) - unconditional logistic regression - assessing the association between each risk factor and breast cancer.

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<56	69	(48.94)	137	(49.28)		
≥56	72	(51.06)	141	(50.72)	.99 (.66-1.48)	.95
Age at first live birth						
<18						
18 to 19	36	(25.53)	56	(20.14)		
20 to 21	36	(25.53)	83	(29.86)	.68 (.38-1.20)	.18
≥22	31	(21.99)	76	(27.34)	.64 (.35-1.15)	.13
Nulliparous	25	(17.73)	52	(18.71)	.75 (.40-1.41)	.37
	13	(9.22)	11	(3.96)	1.83 (.74-4.55)	.19
Age at onset of menstruation						
≥14	39	(27.66)	98	(35.25)		
13	36	(25.53)	69	(24.82)	1.31 (.76-2.27)	.33
12	43	(30.50)	60	(21.58)	1.80 (1.05-3.09)	.03
<12	23	(16.31)	51	(18.35)	1.13 (.61-2.10)	.69
Number of previous benign breast biopsies						
0	111	(78.72)	221	(79.50)		
≥1	30	(21.28)	57	(20.25)	1.05 (.64-1.72)	.85
Total number of first degree relatives with breast cancer						
0						
≥1	104	(73.76)	216	(77.70)		
	37	(26.24)	62	(22.30)	1.24 (.78-1.98)	.37
BMI, kg/m^2						
<25	15	(10.64)	29	(10.43)		
25 to 29.99	39	(27.66)	77	(27.70)	.98 (.47-2.04)	.96
30 to 32.49	34	(24.11)	66	(23.74)	1.00 (.47-2.10)	.99
32.50 to 34.99	22	(15.60)	29	(10.43)	1.47 (.64-3.38)	.37
≥35	31	(21.99)	77	(27.70)	.78 (.37-1.65)	.51
Parity (# of live births)						
≥5						
3 to 4	47	(33.33)	103	(37.05)		
1 to 2	46	(32.62)	108	(38.85)	.93 (.57-1.52)	.78
Nulliparous	35	(24.82)	56	(20.14)	1.37 (.79-2.36)	.26
	13	(9.22)	11	(3.96)	2.59 (1.08-.6.20)	.03

Table 7. Model 4 - Phase II analysis (new categories) - multivariate unconditional logistic regression model – includes risk factors included in the Gail model

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<56	69	(48.94)	137	(49.28)		
≥56	72	(51.06)	141	(50.72)	.98 (.64-1.51)	.94
Age at first live birth						
<18						
18 to 19	36	(25.53)	56	(20.14)		
20 to 21	36	(25.53)	83	(29.86)	.70 (.39-1.26)	.23
≥22	31	(21.99)	76	(27.34)	.63 (.34-1.15)	.13
Nulliparous	25	(17.73)	52	(18.71)	.75 (.39-1.43)	.38
	13	(9.22)	11	(3.96)	1.99 (.79-5.02)	.15
Age at onset of menstruation						
≥14	39	(27.66)	98	(35.25)		
13	36	(25.53)	69	(24.82)	1.35 (.77-2.38)	.29
12	43	(30.50)	60	(21.58)	1.81 (1.05-3.14)	.03
<12	23	(16.31)	51	(18.35)	1.05 (.55-2.00)	.88
Number of previous benign breast biopsies						
0	111	(78.72)	221	(79.50)		
≥1	30	(21.28)	57	(20.25)	1.08 (.64-1.81)	.78
Total number of first degree relatives with breast cancer						
0						
≥1	104	(73.76)	216	(77.70)		
	37	(26.24)	62	(22.30)	1.29 (.80-2.09)	.30

Table 8. Model 5 - Phase II analysis (new categories) - multivariate unconditional logistic regression model – includes risk factors included in the Gail model plus BMI

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<56	69	(48.94)	137	(49.28)		
≥56	72	(51.06)	141	(50.72)	.98 (.64-1.52)	.94
Age at first live birth						
<18						
18 to 19	36	(25.53)	56	(20.14)		
20 to 21	36	(25.53)	83	(29.86)	.70 (.39-1.26)	.23
≥22	31	(21.99)	76	(27.34)	.63 (.34-1.16)	.14
Nulliparous	25	(17.73)	52	(18.71)	.76 (.40-1.47)	.42
	13	(9.22)	11	(3.96)	1.97 (.77-5.00)	.16
Age at onset of menstruation						
≥14	39	(27.66)	98	(35.25)		
13	36	(25.53)	69	(24.82)	1.40 (.79-2.47)	.25
12	43	(30.50)	60	(21.58)	1.89 (1.08-3.31)	.03
<12	23	(16.31)	51	(18.35)	1.10 (.57-2.11)	.77
Number of previous benign breast biopsies						
0	111	(78.72)	221	(79.50)		
≥1	30	(21.28)	57	(20.25)	1.06 (.63-1.79)	.83
Total number of first degree relatives with breast cancer						
0						
≥1	104	(73.76)	216	(77.70)		
	37	(26.24)	62	(22.30)	1.28 (.79-2.08)	.31
BMI, kg/m^2						
<25	15	(10.64)	29	(10.43)		
25 to 29.99	39	(27.66)	77	(27.70)	1.04 (.49-2.19)	.93
30 to 32.49	34	(24.11)	66	(23.74)	.97 (.45-2.10)	.95
32.50 to 34.99	22	(15.60)	29	(10.43)	1.44 (.61-3.41)	.41
≥35	31	(21.99)	77	(27.70)	.76 (.35-1.64)	.48

Table 9. Model 6 - Phase II analysis (new categories) - multivariate unconditional logistic regression model (includes all study risk factors except age at first live birth)

	Cases		Controls			
	N	%	N	%	OR (95% CI)	*P*
Age at screening						
<56	69	(48.94)	137	(49.28)		
≥56	72	(51.06)	141	(50.72)	1.03 (.64-1.65)	.91
Age at onset of menstruation						
≥14	39	(27.66)	98	(35.25)		
13	36	(25.53)	69	(24.82)	1.44 (.82-2.54)	.21
12	43	(30.50)	60	(21.58)	1.93 (1.10-3.38)	.02
<12	23	(16.31)	51	(18.35)	1.20 (.63-2.27)	.58
Number of previous benign breast biopsies						
0	111	(78.72)	221	(79.50)		
≥1	30	(21.28)	57	(20.25)	1.05 (.62-1.77)	.87
Total number of first degree relatives with breast cancer						
0						
≥1	104	(73.76)	216	(77.70)		
	37	(26.24)	62	(22.30)	1.27 (.79-2.06)	.33
BMI, kg/m^2						
<25	15	(10.64)	29	(10.43)		
25 to 29.99	39	(27.66)	77	(27.70)	1.01 (.48-2.13)	.99
30 to 32.49	34	(24.11)	66	(23.74)	.91 (.42-1.96)	.81
32.50 to 34.99	22	(15.60)	29	(10.43)	1.46 (.62-3.44)	.39
≥35	31	(21.99)	77	(27.70)	.73 (.34-1.58)	.42
Parity (# of live births)						
≥5						
3-4	47	(33.33)	103	(37.05)		
1-2	46	(32.62)	108	(38.85)	1.00 (.59-1.70)	.99
Nulliparous	35	(24.82)	56	(20.14)	1.47 (.81-2.66)	.20
	13	(9.22)	11	(3.96)	2.87 (1.16-.7.12)	.02

CHAPTER 5

V. DISCUSSION

This study focused on 141 cases and 278 controls. Aside from nulliparity, I did not find an association for American Indian women in North Dakota between most of the risk factors commonly identified in other populations and breast cancer. The risk factors that were expected to be positively associated with breast cancer, for the most part, showed a null and/or an inverse relationship. The c-statistic for all models in Phase I and Phase II of the analysis ranged from .59-.63. This statistic indicates poor model discrimination meaning that the models don't contain variables that are strongly associated with the outcome of breast cancer. If the regression models contained explanatory variables that were strongly associated with the outcome, improved discrimination is possible.[28] These results could have occurred for a number of reasons, which I will attempt to address below.

A. Chance

My study findings may be due simply to chance (random error). The sample size was small so chance is a possibility. Even though the upper confidence intervals I observed are consistent with point estimates observed in studies conducted with other minority populations[10–12], a larger sample size may have been more informative and would have also allowed for a stratified analysis, for example, by menopausal status.

B. Bias

Potential biases that need to be considered in my study include selection bias, information bias and confounding.[29] My study was hospital-based, not population based

and subjects were required to have had a mammogram. Selection bias may occur because women who choose to get mammograms may differ from women who do not. American Indian women who have been screened for breast cancer via mammogram at Indian Health Service may not be representative of a wider population of American Indian woman because high risk women are more likely to be referred to screening.[14] It may be with a population-based study, findings would be more consistent with what was found in previous studies. Of more concern is the potential for information bias. Since data was incomplete and possibly inaccurate at the Indian Health Service facility, spurious associations may be introduced if the incompleteness or inaccuracy affects those with breast cancer and those without to an unequal degree.

Missing data can also lead to the statistical analysis being biased when more than 10% of the date are missing. Less than 10% of my data was missing for all variables. Data could have been missing for a number of reasons. It may be that the patient opted not to provide the requested information, or there were administrative oversights in the collection, or recorded information did not have enough detail. I dealt with missing data in the analysis by imputing the missing values. Given the small percentage of missing data and data not appearing to be differentially missing between cases and controls, it is unlikely that missing data and imputation led to differential misclassification.

Due to lack of randomization, it is difficult to control for confounding. Therefore, it is important to explore additional risk factors as potential confounders. Numerous risk factors have been identified in recent years (see below) and have yet to be included in a breast cancer risk prediction model.

C. The Gail Model in Other Racial/Ethnic Groups

Risk factors and the magnitude of those risk factors (Gail model risk factors, hormone replacement therapy, alcohol use, body size, oral contraceptive use, positive history of benign breast disease, obesity and lower breast feeding) have varied in other minority populations. Latino women have a greater parity which is believed to contribute, in part, to the lower breast cancer rate compared to Non-Hispanic White women.[30,31] Other factors believed to contribute to the lower breast cancer rate are their younger age at first birth, later age at menarche, lower frequency of use of hormone replacement therapy and very low alcohol use.[31] These factors however typically explain less of the breast cancer risk among Hispanic women compared to Non-Hispanic Whites.[32,33] When exploring body size and breast cancer risk in Hispanic and African American women, the results have been inconsistent, with studies reporting positive, inverse and null results.[33–36] Some studies with African American women show risk associated with early age at menarche, late age at first term pregnancy, nulliparity, older age at menopause, oral contraceptive use, positive family history of breast cancer, positive history of benign breast disease and obesity.[37] African American women with a family history of breast cancer and previous benign breast biopsies are at increased risk of breast cancer compared to other African American women; however these women are at decreased breast cancer risk compared to Non-Hispanic Whites.[12] Another study found that young African American women with lower breast feeding, elevated waist-to-hip ratio and higher parity were at increased risk of breast cancer.[38] Although most risk factors seem to be similar among African American women, considerable variation exists among studies.[12,37] Many studies have an insufficient number of African American

women. As a result, firm conclusions cannot be made as to whether or not risk estimates differ between African American and White women.

It is important to note that comparisons across studies can be challenging because of differences in the analytic approach. Risk factors utilized in the Gail model have been identified as risk factors in other minority populations (see Appendix 1, Table 2). However, these factors do not account for all the differences seen in breast cancer rates between populations. Differences unaccounted for are most likely due to risk factors that have yet to be considered.

D. Additional Risk Factors to Consider in American Indian Women

Risk factors for American Indian women in the Great Plains region may differ and/or vary in magnitude compared to those identified in other populations. Additional breast cancer risk factors have been identified over the past two decades but have not been included in any models.[15] These risk factors include oral contraceptive use, alcohol drinking, active smoking, obesity, diabetes mellitus, menopausal status and breastfeeding. Several of these recently identified risk factors likely occur more frequently in American Indian than other communities.

E. Other Limitations

The type of study and analysis conducted also affects study findings. It may be that results would vary if data were analyzed differently. Various studies have stratified data by looking at risk factors specific to premenopausal women and/or postmenopausal women.[33,39,40] A number of factors that increase the risk of breast cancer have been identified in postmenopausal women. These factors include having a twin birth, increased birth weight, childhood obesity, adult height, obesity,[41] young age at menarche,

age at first live birth >30 years, nulliparity, age at menopause (risk -3% for each additional year) and hormone therapy.[39] Data can also be stratified by breast cancer type. Various subtypes might have different risk factors when modeled independently that vary in magnitude. In order to analyze data in this way, a study would need an ample sample size to maximize the power of detecting a statistically significant comparison. I was unable to carry out such analyses due to the small number of breast cancer cases, and because I did not have information on breast cancer subtype.

It is unlikely though that a single Breast Cancer Risk Assessment Risk Tool could be developed that would work for all tribes or regions of American Indian women. Risk factors for American Indian women in the Great Plains region likely differ from risk factors for American Indians located in other regions. A total of 564 tribes are federally recognized across the United States, each one differing in their culture, environment, healthcare and behavioral health. As a result, the cancer burden can vary significantly from tribe to tribe and these differences could impact the types of risk factors that are present.

F. Conclusions

Currently a model for predicting breast cancer risk does not exist for American Indians. Conducting this study was an important first step in gaining a better understanding of the breast cancer risk factors among American Indian women in Belcourt, ND. The results support the need to explore additional potential breast cancer risk factors in this population, such as oral contraceptive use, alcohol drinking, active smoking, obesity, diabetes mellitus, menopausal status and breastfeeding. ***Once risk factors are identified in American Indian communities, a tool can then be developed***

that will allow American Indian women to find out what their risk of breast cancer is and what factors increase their risk of breast cancer so that appropriate interventions can be designed and implemented at individual, community and policy levels.

At the individual level, a model for predicting breast cancer risk can be used by clinicians as a tool to educate themselves not only about the breast cancer risk factors specific to the American Indian population in their community but about the relative risk of breast cancer for women in the American Indian community given a certain breast cancer risk profile. The model can also be used by clinicians to help them formally weigh the risks and benefits of mammography screening among American Indians. For example, should American Indian woman in her 40s have a screening mammography? The United States Preventive Service Task Force recommends against routine screening mammography in women aged 40 to 49 years. It may be that a model for predicting breast cancer risk would disagree with the standard in the general population and lead the clinician to decide to screen an individual in this population earlier if the model shows that the risk of breast cancer doubles for someone in this age group. Consider the following scenario. A 40-year-old American Indian woman is uncertain whether to have screening mammograms. Her mother and sister had breast cancer. According to the Gail model her risk exceeds that of a 50-year old American Indian woman without risk factors. According to the United States Preventive Service Guidelines, only the 50-year-old American Indian woman should get mammography even though we have every reason to believe that the 40-year-old American Indian woman with risk factors is more likely to benefit from mammography screening than the 50-year-old women who has no risk factors. Having an American Indian model for predicting breast cancer risk allows

clinicians who are advising American Indian women about options for breast cancer prevention to share a general perspective on risk specific to American Indian populations. Also, once the patient knows the risk factors and their own risk, they then have a specific context to consider clinical decisions including whether or not they should be screened. A model for predicting breast cancer risk can also be used by clinicians to determine whether or not to recommend certain secondary prevention treatments, e.g. whether or not to take tamoxifen, given a certain breast cancer risk profile.

At the community level, researchers can conduct randomized controlled trials using a model for predicting breast cancer risk to inform the design of their breast cancer prevention trial. Using risk models allows researchers to figure out how many events are likely to occur over the trial. They can then determine how many women they need to study and for how long they need to study them in order to have a well-designed powerful study. The model can also be used to target a higher risk subset of the population when designing community interventions which allows for more targeted allocation of scarce public health resources.

At the institutional level, American Indian serving facilities will be able to offer their clinicians access to an online breast cancer risk assessment tool so that they are able to have informed conversations with their patients. At the institutional policy level, findings from a model for predicting breast cancer risk can be used to inform screening recommendations in American Indian serving facilities. It may be that these recommendations differ from what is recommended for the general populations by entities such as the United States Preventions Services Task Force.

In summary, it is important to have data driven, population-specific breast cancer risk estimates. By having breast cancer risk estimates, clinicians will be better able to predict breast cancer risk at the individual and population levels. In developing risk estimates, researchers will be able to decide which breast cancer risk factors are important in the American Indian population. Once researchers, clinicians and public health workers are familiar with the breast cancer risk profile for American Indians, they will be better able to decide on appropriate interventions as well as raise awareness of these risk factors in hopes of improving the overall health of the service population.

Disparities in breast cancer incidence, mortality and screening among Northern Plains American Indians emphasize the need to better understand the risk factors associated with breast cancer in this population. Potential next steps include: conducting a population based study with an ample sample size to allow for stratified analyses, exploring breast cancer risk factors in other geographical regions with a substantial American Indian population and exploring additional risk factors for breast cancer in American Indians. It is my hope that this study will contribute to the development of a National Cancer Institute Breast Cancer Risk Assessment Tool that reflects the risk of breast cancer among American Indian women. However, if my results are valid, the future for risk prediction models in American Indian communities is uncertain and clinicians should be cautious in using the current Breast Cancer Risk Assessment Tool to inform their breast cancer risk estimates in the American Indian population.

CHAPTER 6

VI. RESEARCH CHALLENGES AND RECOMMENDATIONS

In the process of conducting my study, I encountered a number of challenges. Although I was able to successfully address issues that arose, I feel it is important to reflect and share my experience. I will also share potential solutions to address each of these challenges.

Estimating breast cancer incidence rates was not the focus of my study; however it was necessary to use published incidence rates to determine the feasibility of the study. IHS site specific data is not available without agency approved agreements between the investigator and the Indian Health Service. The state of North Dakota publishes breast cancer incidence rates; however these rates are published based on county versus tribal land base. Hence it would be ideal for researchers to have access to site specific Indian Health Service data to determine study feasibility. Indian Health Service may want to consider creating and implementing a timely process for researchers to request site-specific prevalence and/or incidence data in order to determine study feasibility

Most studies conducted with non-IHS facilities only need to seek approval from a single IRB. The approval process is quite different when working with sovereign tribal communities. In order to conduct my study, I had to receive approval from the following entities: (1) Indian Health Service Great Plains IRB, (2) University of Minnesota IRB, (3) Turtle Mountain Band of Chippewa and (4) the Indian Health Service Clinic CEO in Belcourt, ND. In addition, any presentation of study results needed to be approved by the Aberdeen area IRB and the Turtle Mountain Band of Chippewa Indians Research Review Board. When I initially approached the Indian Health Service Institutional Review Board

to seek study approval, I was informed that I would need a "letter of support" from the CEO of the Indian Health Service Clinic in Belcourt, ND in order to have my study approved. The CEO denied my request. Indian Health Service is not able to take a stand as to whether or not a study should be conducted or advocate for a study to be funded. What was provided by the CEO was a letter of accommodation. By providing a letter of accommodation, the CEO agrees that the study will not unduly encumber IHS resources, either in personnel time or lab reagents, copy paper etc. Hence, researchers should request a letter of accommodation versus a letter of support. Researchers also need to ensure that they receive approval from proper entities prior to sharing any study findings.

In order to conduct research at the Indian Health Service, I was required to complete a background check as well as HIPPA, Privacy Act, and Information System Security trainings. The background check and trainings can take up to a month to complete. I was also expected to know how to utilize the Resource and Patient Management System. Having previously worked within the Indian Health Service, this was not an issue for me. However, this could pose an issue for a researcher that is not familiar with this system. This is the query software that is used to analyze electronic medical records. Researchers should allocate adequate time to allow for completion of background check and relevant training prior to data collection.

After my background check was cleared and I completed the trainings, I was given card access to the secured area where medical records are stored and was given permission to begin my study. I was working full-time in another town and had to travel 4 hours in order to collect data. I averaged one data collection day per week. The approval is only good for 6 months before having to repeat this process. Fortunately, the

6 months can be based on days at the site collecting data versus consecutive days. However IHS human resources will need to seek approval for time to accrue in this fashion. Researchers should make sure that their time is approved according to study collection days versus consecutive days.

Personal computers are not allowed in secured areas at the facility unless the computer is cleared. Currently it can take up to 2 years to get personal computer clearance. If researchers do not have their computer cleared, they will need to hand write all data collected. Indian Health Service may want to consider creating a process for researchers that shortens this approval process or seek alternative ways to support efficient electronic data collection such as the use of a computer in house and a portable drive that the researcher can use at the facility and on their personal computer off site when needed.

Once I started my study, I wasn't designated a computer to use for data collection. There are computers designated for physicians to use in general which means the researcher will likely utilize a different computer each time data is collected. This can be problematic when trying to keep records on site, such as data collected, the grant, approvals, and training certifications. Indian Health service may want to consider having a space designated for researchers. Medical records staff does assist with record pulls, querying the system, and requesting archived charts; however, these duties are secondary to their primary duties. This issue could possibly be mitigated by training and allowing the researcher to perform these duties. I was able to mitigate the lack of consistent staff support by having a clinical mentor guide me throughout the process. My mentor was previously a provider in the facility and was familiar with the electronic system as well as

clinical processes. My mentor was available to me because I had worked for him previously. Others may not have access to such an individual. Researchers should be mindful that flexibility in gathering records may also be necessary as to not get in the way of department employees.

Once deceased, a patient's radiology file becomes inactive and is refiled by year of death and destroyed after 7 years. This became an issue during my data collection because not all individuals with breast cancer had a mammogram sheet to refer to for data collection. Information found on the mammogram form is not information that is systematically recorded at the clinical level. If a patient's physical chart file becomes inactive after 7 years (they are deceased or no longer seeking care at the facility), the file is sent to the Federal Records Repository in Denver, CO. Instead of destroying the mammogram sheet along with the radiology file, the sheet should be pulled from the radiology file and filed in the medical chart prior to the medical chart being sent to the repository in Denver, CO.

Records that are in the repository are only requested 5 at a time and it can take up to a month between record requests. The medical records department is responsible for paying the shipping and handling for all records requested. The department does not have funding to financially support a record request of this magnitude and as a federal entity, cannot accept payment from an outside source. I worked with a mentor to secure the funding needed to support the record request. My mentor contacted the director of the Indian Health Service who put him in contact with the Indian Health Service Headquarters Office of Public Health Support. Staff at Indian Health Service headquarters drafted a request for senior staff to approve transfer of funds from the Office

of Public Health Support to the service unit in Belcourt, ND. Once appropriate signatures were obtained, approval to proceed with the transfer of funds was granted. A total of $1070 was transferred to cover the cost of the records. Given the lack of funding internally available to support the request of records, the facility may want to explore avenues to accept outside funding to support records request and other pertinent research activities. If a researcher will be requesting archived records from Denver, CO, they should also allocate enough time for data collection.

In summary, I dealt with a number of unforeseen issues throughout the data collection process. However, processes can be implemented at the institutional level to mitigate these issues to further support and encourage researchers use of IHS data. Researchers can also work towards mitigating these issues by having proper foresight of potential issues that may arise.

REFERENCES

1. Breast Cancer Risk Assessment Tool. *National Cancer Institute* (2011). at <http://www.cancer.gov/bcrisktool/>
2. Gail, M. H. *et al.* Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J. Natl. Cancer Inst.* **81,** 1879–1886 (1989).
3. Tice Steven Ziv,Elad Kerlikowske, Karla, J. C. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. *Breast Cancer Res. Treat.* **94,** 115–122 (2005).
4. Tice Rei Adduci, Kelly Petrakis, Nicholas King,Eileen Wrensch, Margaret, J. M. Nipple aspirate fluid cytology and the Gail model for breast cancer risk assessment in a screening population. *Cancer Epidemiol. biomarkers Prev.* **14,** 324–328 (2005).
5. Chen David Ayyagari, Rajeev Graubard, Barry Schairer, Catherine Byrne, Celia Benichou,Jacques Gail, Mitchell, J. P. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *J. Natl. Cancer Inst.* **98,** 1215–1226 (2006).
6. Decarli Stefano Masala, Giovanna Specchia, Claudia Palli,Domenico Gail, Mitchell, A. C. Gail model for prediction of absolute risk of invasive breast cancer: independent evaluation in the Florence-European Prospective Investigation Into Cancer and Nutrition cohort. *J. Natl. Cancer Inst.* **98,** 1686–1693 (2006).
7. Novotny, J. *et al.* Breast cancer risk assessment in the Czech female population--an adjustment of the original Gail model. *Breast Cancer Res. Treat.* **95,** 29–35 (2006).
8. Tice Steven Smith Bindman, Rebecca Ichikawa, Laura Barlow,William Kerlikowske, Karla, J. C. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann. Intern. Med.* **148,** 337–347 (2008).
9. Pastor Barriuso Nieves Ederra, María Erdozáin, Nieves Murillo, Alberto Alés Martínez,José Pollán, Marina, R. A. *et al.* Recalibration of the Gail model for predicting invasive breast cancer risk in Spanish women: a population-based cohort study. *Breast Cancer Res. Treat.* **138,** 249–259 (2013).
10. Banegas, M. P. *et al.* Evaluating breast cancer risk projections for Hispanic women. *Breast Cancer Res. Treat.* **132,** 347–353 (2012).
11. Matsuno, R. K. *et al.* Projecting individualized absolute invasive breast cancer risk in Asian and Pacific Islander American women. *J. Natl. Cancer Inst.* **103,** 951–61 (2011).
12. Gail, M. H. *et al.* Projecting Individualized Absolute Invasive Breast Cancer Risk in African American women. *J. Natl. Cancer Inst.* **99,** 1782–1792 (2007).
13. White, M. C. *et al.* Disparities in cancer mortality and incidence among American Indians and Alaska Natives in the United States. *Am. J. Public Health* **104,** S377-

87 (2014).

14. Chlebowski, R. T. *et al.* Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J. Natl. Cancer Inst.* **97,** 439–48 (2005).
15. Anothaisintawee, T., Teerawattananon, Y., Wiratkapun, C., Kasamesup, V. & Thakkinstian, A. Risk prediction models of breast cancer: a systematic review of model performances. *Breast Cancer Res. Treat.* **133,** 1–10 (2012).
16. Gail, M. Using Risk Models for Breast Cancer Prevention. (2013). at <http://www.youtube.com/watch?v=IDifkvm58hA>
17. Jim, M. a *et al.* Racial Misclassification of American Indians and Alaska Natives by Indian Health Service Contract Health Service Delivery Area. *Am. J. Public Health* **104,** S295–S302 (2014).
18. Howlader, N. *et al. SEER Cancer Statistics Review, 1975-2010*. (2013). at <http://seer.cancer.gov/csr/1975_2010/>
19. Norris, T., Vines, P. L. & Hoeffel, E. M. The American Indian and Alaska Native Population: 2010. *United States Census Bur.* **January,** 1–21 (2012).
20. Wiggins, C. L. *et al.* Cancer among American Indians and Alaska Natives in the United States, 1999-2004. *Cancer* **113,** 1142–52 (2008).
21. Espey, D. K. *et al.* Methods for Improving the Quality and Completeness of Mortality Data for American Indians and Alaska Natives. *Am. J. Public Health* 1–9 (2014). doi:10.2105/AJPH.2013.301716
22. Swan, J. & Edwards, B. Cancer rates among American Indians and Alaska Natives: is there a national perspective. *Cancer* **98,** 1262–1272 (2003).
23. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer Statistics , 2016. *CA. Cancer J. Clin.* **66,** 7–30 (2016).
24. White, A., Richardson, L. C., Li, C., Ekwueme, D. U. & Kaur, J. S. Breast Cancer Mortality Among American Indian and Alaska Native Women, 1990-2009. *Am. J. Public Health* 1–7 (2014). doi:10.2105/AJPH.2013.301720
25. North Dakota Statewide Cancer Registry. (2016). at <https://ndcancer.org/index.html>
26. North Dakota Statewide Cancer Registry. (2016).
27. US Department of the Interior. Great Plains Region. (2016). at <http://www.bia.gov/WhoWeAre/RegionalOffices/GreatPlains/WeAre/Agencies/index.htm>
28. Steyerberg, E. W. *et al.* Assessing the Performance of Prediction Models. *Epidemiology* **21,** 128–138 (2010).
29. Dorak, M. T. Bias & Confounding. **2013,** (2012).
30. Kolonel, L. N., Altshuler, D. & Henderson, B. E. THE MULTIETHNIC COHORT STUDY: EXPLORING GENES, LIFESTYLE AND CANCER RISK. *Nature* **4,** 1–9 (2004).
31. Pike, M. C. *et al.* Breast Cancer in a Multiethnic Cohort in Hawaii and Los

Angeles: Risk Factor-adjusted Incidence in Japanese Equals and in Hawaiians Exceeds that in Whites. *Cancer Epidemiol. Biomarkers Prev.* **11,** 795–800 (2002).

32. Haile, R. W. *et al.* A Review of Cancer in U. S. Hispanic Populations. *Cancer Prev. Res.* **5,** 150–164 (2012).
33. Hines, L. *et al.* Comparative analysis of breast cancer risk factors among Hispanic and non-Hispanic white women. *Cancer* **116,** 3215–3223 (2010).
34. Sexton, K. R. *et al.* A Review of Body Size and Breast Cancer Risk in Hispanic and African American Women. *Cancer* **117,** 5271–5281 (2011).
35. Bandera, E. V, Maskarinec, G., Romieu, I. & John, E. M. Racial and Ethnic Disparities in the Impact of Obesity on Breast Cancer Risk and Survival: A Global Perspective. *Adv. Nutr.* **6,** 803–819 (2015).
36. Adderly-Kelly, B. & Williams-Stephens, E. The relationship Between obesity and breast cancer. *ABNF J.* **14,** 61–65 (2003).
37. Bernstein, L., Teal, C. R., Joslyn, S. & Wilson, J. Ethnicity-Related Variation in Breast Cancer Risk Factors. *Cancer* **97,** 222–229 (2003).
38. Danforth, D. J. Disparities in breast cancer outcomes between Caucasian and African American women: a model for describing the relationship of biological and nonbiological factors. *Breast Cancer Res.* **15,** 11–16 (2013).
39. Vogel, V. G. Epidemiology , genetics , and risk evaluation of postmenopausal women at risk of breast cancer. *Menopause* **15,** 782–789 (2008).
40. Velie, E. M., Nechuta, S. & Osuch, J. R. Lifetime Reproductive and Anthropometric Risk Factors for Breast Cancer in Postmenopausal Women. *Breast Dis.* **24,** 17–35
41. Stephenson, G. D. & Rose, D. P. Breast Cancer and Obesity: An Update. *Nutr. Cancer* **45,** 1–16 (2016).

APPENDIX 1

Table 1. Select baseline characteristics and breast cancer risk factors by race/ethnicity[14]

	Whites (n=129037)		American Indians/Alaska Natives (n=696)	
Characteristic	n	%	n	%
Number of first-degree relatives with breast cancer				
0	102454	79.4	527	75.7
1	16479	12.8	92	13.2
≥2	1697	1.3	8	1.2
Age of onset of menstruation,				
<12				
12-13	27732	21.5	167	24.0
≥14	71981	55.8	328	47.1
	28867	22.4	195	28.0
Age at first birth,				
Never giving birth	14868	11.5	51	7.3
<20	14589	11.3	173	24.9
20-24	50972	39.5	230	33.1
25-29	28698	22.2	80	11.5
≥30	9460	7.3	45	6.5
Number of previous benign breast biopsies	26673	20.6	127	18.2
BMI, kg/m^2				
<25	47222	36.6	165	23.7
25-30	44729	34.7	207	29.7
≥30	35982	27.9	309	44.4
Number of full-term pregnancies				
Nulliparous	14838	11.5	50	7.2
1	10368	8.0	81	11.6
2	32649	25.3	131	18.8
3	32096	24.9	164	23.6
4	20121	15.6	104	14.9
≥5	18254	14.2	160	23.0

Table 2. Characteristics of Gail and Modified Gail Studies

Author	Year	Study Design	Ethnicity	Statistical Method	Breast Cancers	Non Breast Cancers	Risk Factors Included in the Model
Gail [2]	1989	Nested case-control	White	Logistic regression	2,852	3,146	Age, age at menarche, age at first live birth, number of first degree relatives with breast cancer, number of previous breast biopsies
Tice[3]	2005	Cohort	Mixed (White, Black, Asian, Latina, Other)	Cox regression	955	80,822	Risk factors in Gail 1989 and breast density
Tice[4]	2005	Cohort	Mixed (White, Black, Asian Latina, Other)	Cox regression	400	6,504	Risk factors in Gail 1989 and nipple aspirate fluid cytology
Chen[5]	2006	Nested case-control	White	Logistic regression	1,280	4,035	Age, age at birth of first child, family history of breast cancer, numbers of previous breast biopsy, breast density, weight
Decarli[6]	2006	Case-control	Italian (conducted in Italy)	Logistic regression	2,569	2,588	Risk factors in Gail 1989
Novotny[7]	2006	Case-control	Mixed (conducted in Czech Republic)	Logistic regression	2,299	2,299	Risk factors in Gail 1989, family history of any cancer, history of breast inflammation, body mass index, parity
Author	Year	Study Design	Ethnicity	Statistical Method	Breast Cancers	Non Breast Cancers	Risk Factors Included in the Model

Table 2 cont. Characteristics of Gail and Modified Gail Studies

Gail[12]	**2007**	**Case-control**	**African American**	**Logistic regression**	**1,607**	**1,647**	**Risk factors in Gail 1989**
Tice[8]	2008	Cohort	Mixed (White, Black, Asian, Hispanic, Other)	Cox regression	14,766	629,229	Age, race, breast density, family history of breast cancer, history of breast biopsy
Matsuno[11]	2011	Case-control	Chinese, Japanese, Filipino, Hawaiian, Other Pacific Islanders, and other Asian women	Logistic regression	589	952	Age at menarche, number of affected mothers, sisters, daughters, number of previous benign breast biopsies
Pastor-Barriuso[9]	2013	Cohort	Spanish (conducted in Spain)	Logistic regression	835	53,814	Risk factors in Gail model

APPENDIX 2

University of Minnesota IRB Approval

UNIVERSITY OF MINNESOTA

Twin Cities Campus

Human Research Protection Program
Office of the Vice President for Research

D528 Mayo Memorial Building
420 Delaware Street S.E.
MMC 820
Minneapolis, MN 55455

Office: 612-626-5654
Fax: 612-626-6061
E-mail: irb@umn.edu or ibc@umn.edu
Website: http://research.umn.edu/subjects/

February 28, 2014

Melanie A Nadeau
1300 S 2nd St
Suite 300
Minneapolis, MN 55454

RE: "Exploring Breast Cancer Risk Factors among American Indian Women"
IRB Code Number: **1402M47864**

Dear Dr. Nadeau:

Medical Record Chart Review; Expedited Review Approval

Message: The Institutional Review Board (IRB) reviewed the referenced study by expedited review procedures and has granted approval under federal guidelines 45 CFR Part 46.110 category (5) Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for non-research purposes (such as medical treatment or diagnosis).

The code number above is assigned to your research. That number and the title of your study must be used in all communication with the IRB office.

This e-mail confirmation is your official University of Minnesota RSPP notification of approval. You will not receive a hard copy or letter. This secure electronic notification between password protected authentications has been deemed by the University of Minnesota to constitute a legal signature.

The date of approval is February 18, 2014. Your study will expire one year from this approval date. A report form will be sent out two months before the expiration date.

You are approved to access 450 records from Quentin N. Burdick Health Care Facility Indian Health Service. Please note that you may be required to present this letter when requesting access to records.

If you requested a waiver of HIPAA Authorization and received this e-mail, the waiver was granted. Please note that under a waiver of the HIPAA Authorization, the HIPAA regulation [164.528] states that the subject has the right to request and receive an accounting of Disclosures of PHI made by the covered entity in the six years prior to the date on which the accounting is requested.

If you are accessing a limited Data Set and received this email, receipt of the Data Use Agreement is acknowledged.

Driven to Discover℠

As the Principal Investigator of this project, you are required by federal regulations to inform the IRB of any proposed changes in your research that will affect human subjects. Changes should not be initiated until IRB approval is received. Unanticipated problems and adverse events should be reported to the IRB as they occur. Research projects are subject to continuing review and renewal. Once data analysis is complete, you must notify the IRB to inactivate your research project. If you have any questions, call the IRB office at 612-626-5654.

Upon receipt of this email, you may begin your research. If you have questions, please call the IRB office at 612-626-5654.

You may go to the View Completed section of eResearch Central at http://eresearch.umn.edu/ to view further details on your study.

The IRB wishes you success with this research.

Sincerely,

Christina Dobrovolny, CIP
Research Compliance Supervisor
CD/bs

CC: Jean Forster

Great Plains IRB approval letter

GREAT PLAINS IRB/RESEARCH and PUBLICATION COMMITTEE

Great Plains IRB
Indian Health Service
Federal Building, Room 309
115 – 4th Ave. SE
Aberdeen, SD 57401
Toll Free #: (866) 331-5794

April 9, 2014

Melanie Nadeau, MPH
PhD in Epidemiology Student
2441 Larpenteur Ave. W, Apt 2
Lauderdale, MN 55113

GPIRB #: 14-R-05GP

Dear Ms. Nadeau,

The Great Plains Institutional Review Board (GPIRB) reviewed the information you submitted for protocol 14-R-05GP "Exploring Breast Cancer Risk Factors among American Indian Women."

This is to confirm that your protocol is approved. Please be aware that according to the publication approval process, *the AAIRB has final approval of ALL publications related to this project*. You are granted permission to conduct your study, as described, effective immediately. The study is subject to continuing review on or before **February 28, 2015**, unless closed before that date.

Please note that any changes to the study as approved must be promptly reported and approved. Some changes may be approved by expedited review; others require full board review. This approval does *not* cover any fliers or presentations (oral, poster, or handouts) that may be made regarding this study. Presentations will need a separate AAIRB approval and Service Unit/Tribal approvals if necessary. Contact Marsha Stevens, AAIRB Coordinator, at (605) 226-7493 if you have any questions or require further information.

Sincerely,

Marsha Stevens
for
Dewey J. Ertz, EdD
Chairman, GPIRB

Dewey Ertz, EdD
Chairman, GPIRB
Psychological Evaluation, Therapy and Research, Ltd.
403 National Street #1
Rapid City, SD 57702
Phone #: (605) 341-8647
FAX #: (605) 341-0489
Email: dertz.chrysalis@midconetwork.com

Marsha Stevens
GPIRB Coordinator
Office of Planning & Legislation
Indian Health Service
Federal Building, Room 309
115 – 4th Ave. SE
Aberdeen, SD 57401
Phone #: (605) 226-7493
FAX #: (605) 226-7214
Email: marsha.stevens@ihs.gov

Letter of Accommodation from Indian Health Service CEO

DEPARTMENT OF HEALTH & HUMAN SERVICES
PUBLIC HEALTH SERVICE
HEALTH SERVICES AND MENTAL HEALTH ADMINISTRATION

Quentin N. Burdick Memorial
Health Care Facility
PHS Indian Hospital
PO Box 160
Belcourt, North Dakota 58316
(701) 477-6111

March 5, 2014

Shelly Harris
1300 Hospital Loop
Belcourt, ND 58316
March 4, 2014

Re: Letter of accommodation

To whom it may concern:

On behalf of the Quentin N. Burdick Memorial Health Care Facility in Belcourt, ND, I am writing a letter of accommodation for Melanie Nadeau's research study entitled "Exploring Breast Cancer Risk Factors among American Indian Women".

I am aware that this study has been approved by both the Aberdeen Area Institutional Review Board and the University of Minnesota Institutional Review Board. In my judgment, the conduct of this study will not significantly divert resources from the provision of health care services to the population served by the facility.

Sincerely,

Shelly Harris
Chief Executive Officer

Turtle Mountain Band of Chippewa Tribal Resolution

RESOLUTION NUMBER **TMBC668-03-14** OF THE DULY ELECTED AND CERTIFIED BODY OF THE TURTLE MOUNTAIN BAND OF CHIPPEWA

WHEREAS, the Turtle Mountain Band of Chippewa, hereinafter referred to as the Tribe, is an unincorporated Band of Indians acting under a revised Constitution and Bylaws approved by the Secretary of the Interior on June 16, 1959, and amendments thereto approved; and

WHEREAS, the Turtle Mountain Constitution and Bylaws was adopted by the tribal citizens to promote the general welfare of tribal citizens, and

WHEREAS, the Tribe is aware of the serious impact of various health conditions on our people; and has traditionally supported efforts of medical researchers and clinicians to improve medical care and prevention of disease; and

WHEREAS, the Tribe understands the importance of gaining full and informed consent for all participants in medical research. In addition, we are committed to assuring confidentiality and privacy for any medical information that research participants provide; and

WHEREAS, the Tribal Nations Research Group recommends approval of the dissertation research presented by Melanie Nadeau titled "Exploring Breast Cancer Risk Factors among American Indian Women".

THEREFORE BE IT RESOLVED that the Tribe hereby approves Melanie Nadeau to perform the dissertation research titled "Exploring Breast Cancer Risk Factors among American Indian Women".

THEREFORE BE IT FURTHER RESOLVED that the researcher understands that the data and/or information collected while doing research can only be used for the purpose of dissertation research, any additional uses need to be approved.

CERTIFICATION

I, the undersigned Tribal Secretary of the Turtle Mountain Band Chippewa Indians, do hereby certify that the Tribal Council is composed of nine (9) members of whom ***eight (8)*** constituting a quorum were present at a meeting duly called, convened and held on the ***27th day of March, 2014,*** that the foregoing resolution was adopted by an affirmative vote of seven (7) in favor-Representatives Edward Falcon, Troy DeCoteau, Cindy Malaterre, Patrick J. Marcellais, Elmer Davis, Jr., Carson Belgarde and Zelma Peltier; one (1) absent-Rep. Jim Baker; with the Chairman not voting.

Jolean A. Morin, Tribal Secretary 3/27/14

(X) SIGNED INTO LAW/Dated this 27th day of March, 2014
( ) VETOED/Dated this ____ day of _____, 2014

Richard W. McCloud, Tribal Chairman

HUMAN SUBJECTS

NIH criteria for being a Phase 3 clinical trial do not apply to this study.

Human Subjects Protection

The Institutional Review Boards of the University of Minnesota and Great Plains Indian Health Service have waived the requirement for written consent of participants for the following reasons: (1) the study is being conducted in the interest of public health, (2) direct harm to participants is extremely unlikely, and (3) individually identifiable data are not collected.

Special Populations

Women. Due to the sex-limited nature of this study, women will be the only participants.

Racial Minorities. Because of the nature of the research question, enrollment will be limited to those of American Indian ethnicity.

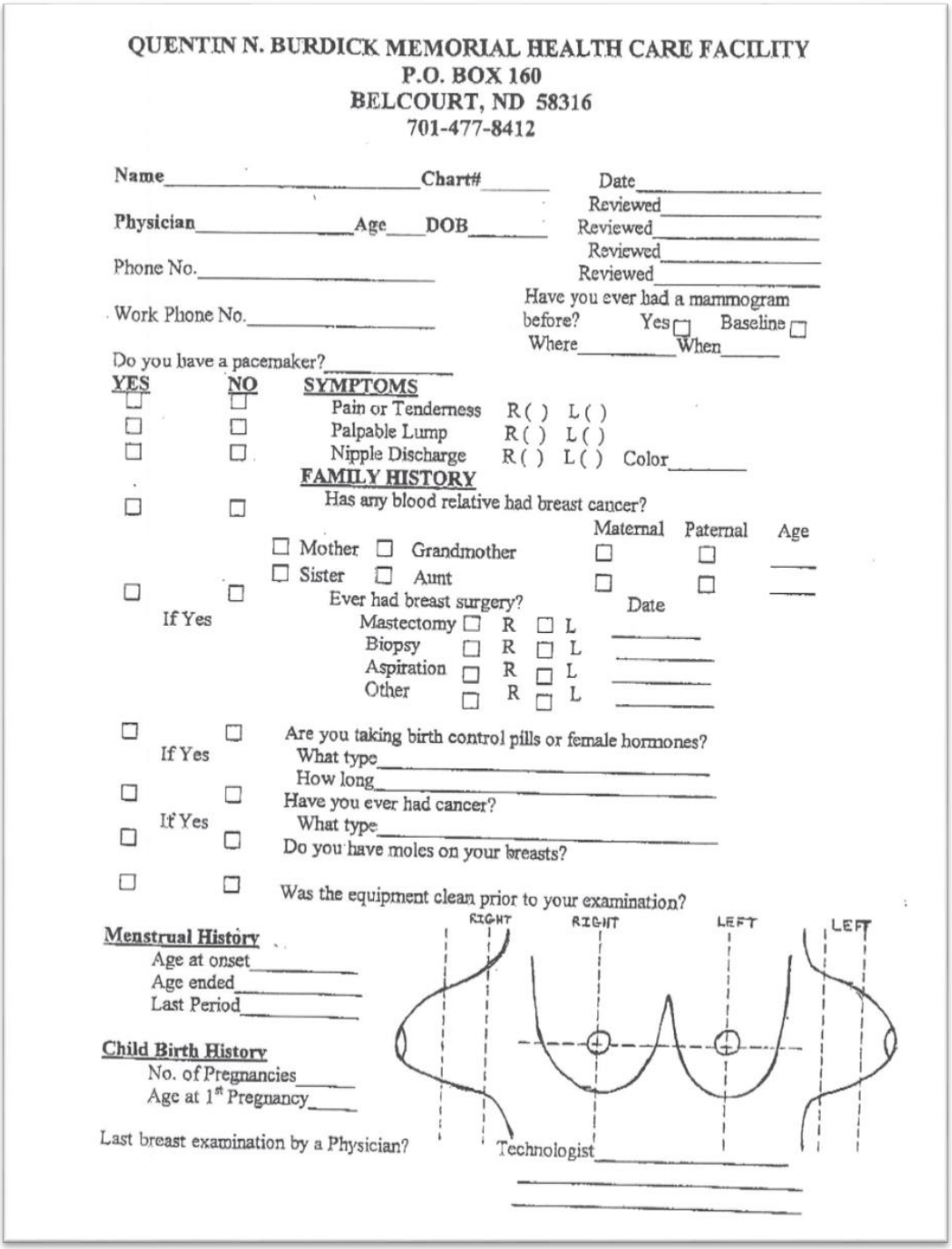
Children. Children will not be included in this study. All participants will be over the age of 18.

Potential Risks

The risk to participants in this study is minimal because I am not collecting private, identifiable data from the medical record.

APPENDIX 3

Example of the questionnaire that is administered at an Indian Health Service facility



QUENTIN N. BURDICK MEMORIAL HEALTH CARE FACILITY
P.O. BOX 160
BELCOURT, ND 58316
701-477-8412

Name_____ **Chart#**_____ Date_____
Reviewed_____
Physician_____ **Age**___ **DOB**_____ Reviewed_____
Reviewed_____
Phone No._____ Reviewed_____
Have you ever had a mammogram
Work Phone No._____ before? Yes ☐ Baseline ☐
Where_____ When_____
Do you have a pacemaker?_____

YES	NO	SYMPTOMS
☐	☐	Pain or Tenderness R( ) L( )
☐	☐	Palpable Lump R( ) L( )
☐	☐	Nipple Discharge R( ) L( ) Color_____

FAMILY HISTORY

☐ ☐ Has any blood relative had breast cancer?

	Maternal	Paternal	Age
☐ Mother ☐ Grandmother	☐	☐	____
☐ Sister ☐ Aunt	☐	☐	____

☐ ☐ Ever had breast surgery?

If Yes				Date
Mastectomy	☐ R	☐ L		_____
Biopsy	☐ R	☐ L		_____
Aspiration	☐ R	☐ L		_____
Other	☐ R	☐ L		_____

☐ ☐ Are you taking birth control pills or female hormones?
If Yes
What type_____
How long_____
☐ ☐ Have you ever had cancer?
If Yes
What type_____
☐ ☐ Do you have moles on your breasts?

☐ ☐ Was the equipment clean prior to your examination?

Menstrual History
Age at onset_____
Age ended_____
Last Period_____

Child Birth History
No. of Pregnancies_____
Age at 1st Pregnancy_____

Last breast examination by a Physician?

Technologist_____

APPENDIX 4

Table 3. Illustration of data collection sheet

Patient ID	1001	1002
Caseness		
Radiology review complete		
Electronic review complete		
Chart archived		
Chart review complete		
Age at screening		
Age at first pregnancy		
Age at first live birth		
Age at onset of menstruation		
Number of previous benign breast biopsies		
Total number of first degree relatives with breast cancer		
BMI before date of diagnosis		
BMI after date of diagnosis		
Number of pregnancies		
Number of live births		