Selective Serotonin Reuptake Inhibitors (SSRIs) and Breast Cancer:

A Record Linkage Study

by

Janet E. Ashbury

A thesis submitted to the Department of Community Health and Epidemiology in conformity with the requirements for the degree of Master of Science

Queen’s University
Kingston, Ontario, Canada
January, 2008

Copyright © Janet E. Ashbury, 2008
Abstract

Evidence suggests that selective serotonin reuptake inhibitors (SSRIs, a class of antidepressant medications) may contribute to increased breast cancer risk by stimulating the secretion of prolactin, a potential tumour promoter. The main objective of this study was to determine breast cancer risk associated with the duration, dosage and timing of SSRI use among women, with control for a limited set of confounders. This thesis project, conducted within the context of a population-based two-stage case-control study, consisted of a record linkage study utilizing three Saskatchewan health services databases. Cases included 1,273 women with primary breast cancer diagnosed between January 1, 2003 and December 31, 2005, and controls consisted of 12,730 subjects randomly selected from the province’s population registry. Data on SSRI use was compiled from the Saskatchewan prescription drug plan database. Information on a limited set of established risk factors for breast cancer that may confound this relationship was ascertained from the population registry and the prescription database.

Cases and controls were similar in terms of age, total number of consecutive years eligible for prescription coverage and indicators of socioeconomic status. Compared to controls, cases were more likely to be married and to have used hormone therapy and/or oral contraceptives.

Compared to nonusers, results indicated that the use of SSRIs for three or more years (as estimated by having filled 36 or more prescriptions for all SSRIs combined during the main exposure window more than two years prior to index date) was not associated with an increased risk of breast cancer (OR= 1.08, 95% CI: 0.74-1.58), controlling for age, marital status, oral contraceptive and hormone therapy use. In
addition, no suggestion of increased risk was detected for long-term exposures to individual SSRIs (24 or more prescriptions filled during the main exposure window) and in relation to total combined SSRI use 2-7 years and more than seven years prior to index date. However, these risk estimates may have been affected by potential sources of information bias and confounding. In summary, these results do not provide evidence to suggest that the risk of breast cancer is increased with the use of SSRIs.
Acknowledgements

I would like to express my sincere, whole-hearted thanks to many individuals who have supported me along the way during this long journey. Most importantly, I wish to express my appreciation and countless thanks to my mentors, supervisors and co-workers, Dr. Kristan Aronson and Dr. Will King for their endless support and encouragement from the day this project idea was suggested more than 5 years ago. I am especially grateful to my primary supervisor Dr. Kristan Aronson, who always believed in me from the beginning and encouraged me to ‘give birth’ to this project after so many years. Her confidence in my abilities, expert advice and guidance, and friendship are most sincerely appreciated. I am also very grateful to Dr. Will King for his day-to-day patience with my requests for help, especially as I stumbled along trying to understand the workings of SAS programming.

I wish to thank many fellow students who offered me their support. Many thanks to Jennifer Goy for her ongoing encouragement and her positive attitude towards my never-ending lists of questions and requests for help solving SAS glitches. I also wish to thank Amy Schneeberg, whose passion for healthy living inspired me to start (and to continue) exercising and whose support along with her computer formatting skills gave me the motivation and help that I needed to put the finishing touches on this project. And I can’t forget to thank John Queenan, my ‘stats buddy’ from first year, for his friendship and upbeat outlook even in the face of many daunting statistic assignments.

I also wish to thank my past and present coworkers at the Cancer Research Unit, including Celine Zakos, Sarah Troster, Maria Sanchez and especially Meghan Hamel who patiently helped me along the way with many aspects of this project and the larger study.
I am also grateful for the support of my friends and coworkers in the Department of Community Health and Epidemiology: Lee Watkins, Katherine Cook, Gwyneth Fairfield, Dana Al-Bargash, and Vikki Ho (for her last minute help with formatting).

I also wish to acknowledge the thoughtful input of Dr. Linda Levesque, who helped me to understand many of the intricacies of pharmacoepidemiology. In addition, I wish to thank Patty Beck from the Saskatchewan Ministry of Health, our co-investigator for the larger project related to this thesis work, for her assistance in explaining many methodological details of this project and her careful editing of the manuscript. Also, I want to acknowledge the help and guidance of Winanne Downey from the Saskatchewan Ministry of Health and Heather Stuart from the Saskatchewan Cancer Agency.

Finally, I would like to express my extreme gratitude to my family, who are my constant inspiration; my children, Andrew, Alie and Jenny, who always provided me with sincere encouragement and warm hugs despite my ongoing preoccupation with finishing this project, and my husband, Ted, who has supported me with constant patience, humour and understanding. It is truly his uncanny ability to always view obstacles from a positive perspective that kept me moving forward with this project.

Disclaimer: This study is based in part on de-identified data provided by the Saskatchewan Ministry of Health. The interpretation and conclusions contained herein do not necessarily represent those of the Government of Saskatchewan or the Saskatchewan Ministry of Health.
# Table of contents

Abstract .................................................................................................................................................. ii  
Acknowledgements ............................................................................................................................ iv  
Table of Contents ................................................................................................................................. vi  
List of Figures and tables .................................................................................................................... ix  
List of Abbreviations ............................................................................................................................. xi

Chapter 1.0 INTRODUCTION .................................................................................................................... 1

Chapter 2.0 LITERATURE REVIEW ......................................................................................................... 2  
2.1 Selective serotonin reuptake inhibitors (SSRIs) ........................................................................... 2  
2.2 Breast cancer .................................................................................................................................. 3  
2.3 SSRIs and breast cancer ................................................................................................................. 7  
  2.3.1 Experimental research ............................................................................................................... 7  
  2.3.2 Epidemiologic research .......................................................................................................... 8  
2.4 Underlying biologic mechanisms .................................................................................................. 15  
  2.4.1 SSRI prolactin hypothesis ....................................................................................................... 15  
  2.4.2 Other SSRI tumour-promoting mechanisms ....................................................................... 17  
2.5 Breast cancer development ......................................................................................................... 18  
  2.5.1 Overview of carcinogenesis and tumour promotion ............................................................... 18  
  2.5.2 Conceptual model of exposure to SSRIs and the development of breast cancer ............... 19  
2.6 Pharmacoepidemiology: investigating the long-term safety of drugs................................. 21  
2.7 Drug exposure measurement issues ............................................................................................. 23  
2.8 Rationale for further study of the relationship between SSRIs and breast cancer ................... 25  
2.9 Context of thesis research ............................................................................................................. 26

Chapter 3.0 METHODS ............................................................................................................................. 28  
3.1 Objectives...................................................................................................................................... 28  
3.2 Overview of thesis research design .............................................................................................. 28  
3.3 Study population ............................................................................................................................ 30  
  3.3.1 Source population ................................................................................................................... 30  
  3.3.2 Summary of case-control sampling process ..................................................................... 31
5.5 Methodological considerations ......................................................... 85
  5.5.1 Strengths .................................................................................. 85
  5.5.2 Limitations .............................................................................. 86

5.6 Summary explanation of null findings ............................................. 95

5.7 Suggestions for future research .................................................... 96

5.8 Summary and contribution of this thesis research ....................... 98

REFERENCES .................................................................................. 99

APPENDICES .................................................................................. 110
  Appendix 1: Research Ethics Board Approval .................................. 110
  Appendix 2: Detailed description of case-control data linkage sampling method ......................................................... 114
  Appendix 3: Prevalence data tables .................................................... 120
  Appendix 4: Alternative confounder assessment ............................ 121
List of Figures and Tables

Figure 1: Initiation and promotion of cancer: .............................................................. 19
Figure 2: Conceptual model of breast cancer development........................................ 21
Figure 3: Overview of subject accrual process for two-stage study ...................... 27
Figure 4: Subject accrual process for thesis project............................................... 35
Figure 5: SSRI exposure windows of interest.......................................................... 43
Figure 6: Potential confounders for SSRI use and breast cancer risk................... 48
Figure 7: Prevalence of SSRI use by year among controls................................. 58
Figure 8: Prevalence of individual SSRI use among SSRI users (controls).......... 60
Figure 9: Frequency distribution of SSRI use among controls............................ 62
Figure 10: Case-control sampling method (detailed)............................................. 119

Table 1: Use of one or more type of SSRIs among controls................................. 60
Table 2: Frequency distribution of SSRI use among controls ............................. 61
Table 3: Characteristics of case and control subjects and unadjusted odds ratios for breast cancer associated with potential confounders.............................. 64
Table 4: SSRI distribution for total SSRI use among cases and controls during exposure windows .............................................................. 65
Table 5: Characteristics related to the use of SSRIs among controls .................. 67
Table 6: Odds ratios for breast cancer according to total number of SSRI prescriptions during each exposure window......................................................... 69
Table 7: Odds ratios for breast cancer according to total SSRI dosage during each exposure window ............................................. 71
Tables 8-12: Odds ratios for breast cancer according to total number of prescriptions for each individual SSRI .......................................................... 72
   Table 8: Paroxetine............................................................................................ 72
   Table 9: Sertraline ....................................................................................... 73
   Table 10: Fluoxetine .................................................................................... 73
   Table 11: Fluvoxamine .............................................................................. 74
   Table 12: Citalopram ................................................................................. 74
Table 13: Odds ratios for breast cancer according to total number of prescriptions for paroxetine, sertraline and/or fluoxetine .......................................................... 75

Tables 14-15: Odds ratios for breast cancer according to total number of SSRI prescriptions during main exposure window by age at diagnosis ......................... 76

Table 14: Subjects 55 years of age and younger ................................................. 76
Table 15: Subjects over age 55 ........................................................................... 77

Tables 16-19: Stage two sampling tables ................................................................ 78

Table 16: Case-control exposure distribution (index date 2004-2005) .......................................................... 78

Table 17: Estimated required number of stage two cases and controls that need to be invited for interviews considering expected response rates ................. 80

Table 18: Actual final sample of cases and controls invited for interviews .... 81

Table 19: Anticipated first stage two sample (2004-2005 index date) with completed interviews .......................................................... 81

Appendix tables .................................................................................................. 120

Table 3.1: Prevalence of SSRI use by year among controls ..................................... 120
Table 3.2: Prevalence of individual SSRI use by year among SSRI users (controls) .................................................................................................................. 120
Table 4.1: Assessment of potential confounders using the change-in-estimate method ........................................................................................................ 121
List of Abbreviations

SSRIs: Selective serotonin reuptake inhibitors
SK: Saskatchewan
OCs: Oral contraceptives
HT: Hormone therapy
SCA: Saskatchewan Cancer Agency
SPDP database: Saskatchewan prescription drug plan database
PRS: Person registry system
HSN: Health services number
ICDO: International classification of diseases for oncology
TNM: Tumour, node, metastasis
DDD: Defined daily dose
SES: Socioeconomic status
Chapter 1.0 Introduction

It is estimated that 5-10% of women take selective serotonin reuptake inhibitors (SSRIs), a group of antidepressant medications. Experimental evidence suggests that SSRIs could promote breast tumour development and therefore contribute to increased breast cancer risk. Few epidemiologic studies have examined this relationship using detailed exposure measurements related to long-term SSRI use. Given projections of increased SSRI use for several medical and psychiatric indications, assessing this biologically plausible association with breast cancer risk is timely.

The objective of this record-linkage nested case-control study was to determine breast cancer risk associated with the dosage, duration and timing of SSRI use among Saskatchewan women, with control for a limited set of confounders. An innovative feature of this study design includes the use of the population-based Saskatchewan prescription database for access to objective, detailed and long-term exposure data related to the use of SSRIs as well as other medications (oral contraceptives and hormone therapy) that could confound the exposure-disease relationship.

This study, which is based on a sub-set of the stage one data currently being collected for a CIHR/CBCRA-funded (Canadian Institutes of Health Research/Canadian Breast Cancer Research Alliance) two-stage project, will provide a framework for conceptualizing SSRI exposures in terms of dosage, duration and critical time periods of exposure and facilitate the final analysis of the complete stage one data. In addition, results of this thesis research will begin to provide a more complete understanding of the possible role of the SSRI group of antidepressants as potentially modifiable risk factors in the complex etiology of breast cancer.
Chapter 2.0 Literature Review

2.1 Selective serotonin reuptake inhibitors (SSRIs)

Selective serotonin reuptake inhibitors (SSRIs) first became available in Canada in the late 1980’s. Fluoxetine (Prozac®) was first introduced in 1989, fluvoxamine (Luvox®) in 1990, sertraline (Zoloft®) in 1992, paroxetine (Paxil®) in 1993, citalopram (Celexa®) in 1999 and escitalopram (Cipralex®) in late 2004. SSRIs are prescribed to treat many psychiatric conditions including depression, and anxiety, obsessive-compulsive and personality disorders. They are also used for the management of a growing number of other conditions such as premenstrual syndrome, chronic pain (fibromyalgia and neuropathic pain), substance use disorders, eating disorders and for migraine headache prophylaxis (Stone et al., 2003; Bahl et al., 2003). Since July 2002, when harmful effects of hormone therapy (HT) were reported, SSRIs have been used increasingly for control of the full range of menopausal symptoms (McIntyre et al., 2005).

SSRIs produce their therapeutic action by blocking the reuptake of the neurotransmitter, serotonin (5-hydroxytryptamine, 5-HT), in presynaptic neurons, resulting in enhanced serotonergic function in the brain and enhanced mood (Ward and Azzaro, 2004). Due to their relative specificity of action, SSRIs offer greater safety overall, fewer side effects and a lower toxicity profile compared to tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) (Ward and Azzaro, 2004; Nelson et al., 1997).

The prevalence of SSRI use increased substantially throughout the 1990s, with overall prevalence in women estimated to be 5–10% in the early 21st century (Patten et al., 2005; Bahl et al., 2003; Steingart et al., 2003) and continuing to rise.
prescription data from two United States cross-sectional health examination surveys, Paulose-Ram et al. (2007) reported a three-fold increase in overall antidepressant use among the non-institutionalized adult population (≥ 17 years of age) between 1988-1994 and 1999-2002 (age-adjusted prevalence figures increasing from 2.5% to 8.1% respectively during these two time periods). SSRI users increased from 28% to 60% of all antidepressant users over this period of time (Paulose-Ram et al., 2007).

In Canada, Patten et al. (2005) using telephone research methods for sampling household residents of Alberta between the ages of 18 and 64, reported that 6.1% of women were SSRI users at the time of the interview. Between 1999 and 2005, the total number of prescriptions for SSRIs in Canada increased by 60%, from 6.5 to 10.4 million (market share of 42.9% of all prescribed antidepressants) (IMS Health, Canadian CompuScript, 2005). Age-specific prevalence figures are not available, although some evidence suggests that antidepressants are more often prescribed to women less than 50 years of age compared to women 50 and over (Sclar et al., 1998; IMS Health, Canadian Disease and Therapeutic Index, 2005).

2.2 Breast cancer

Worldwide, breast cancer is the leading cause of death due to cancer in women. Breast cancer is the most frequently diagnosed cancer in Canadian women, accounting for about 30% of all new cancer cases each year (Canadian Cancer Statistics, 2007). Since 1984 breast cancer incidence rates have increased slightly for women overall, and especially in women over 50, while mortality rates have remained relatively stable (Canadian Cancer Statistics, 2007). One in nine women will develop breast cancer in her lifetime and one in 25 will die from the disease (Public Health Agency of Canada, 2007).
There are several well-established risk factors for breast cancer (Reviewed in Zagrofos et al., 2004; Armstrong et al., 2000; Hulka and Moorman 2001; Biglia et al., 2004; Key et al., 2001). Increased risk of developing breast cancer is associated with increasing age, increasing number of first-degree relatives with breast cancer and mutations of the BRCA1 and BRCA2 genes. Several menstrual and reproductive factors have consistently been shown to be associated with elevated risk including younger age at menarche (<12 years), delayed menopause (>55 years), older age at first full-term pregnancy (>30 years) and nulliparity.

Other factors have been shown to be related to breast cancer risk, although the epidemiologic evidence has been inconsistent (Reviewed in Zagrofos et al., 2004; Hulka and Moorman, 2001; Biglia et al., 2004; Key et al., 2001). The role of nutritional and lifestyle factors in the development of breast cancer has been studied extensively. Studies indicate that higher rates of consumption of vegetables, fruits, fish, oils and dietary fibre may be related to a decrease in risk. On the other hand, higher intake of meat (especially red meat), sweets and saturated fats may be associated with an increased risk. Studies also suggest that a lack of physical activity, increasing alcohol consumption and smoking (earlier onset, longer duration and higher intensity) increase risk. Further, occupational exposures to asbestos, organic solvents and ionizing radiation have been shown to be related to elevated risk. Anthropometric factors associated with elevated risk include increased height, higher body mass index (BMI) and higher waist-to-hip ratio.

Higher socioeconomic status, as measured by higher income, education and occupation level, has been shown to be associated with elevated breast cancer risk. Also, evidence suggests that women who have had a previous breast biopsy reported as atypical ductal or lobular hyperplasia are at a higher risk of developing breast cancer compared to
women without these abnormalities (Arpino et al., 2005). Increased mammographic breast density is a strong risk factor for breast cancer and has been consistently associated with other hormonally-related risk factors such as younger age, pre-menopausal status, nulliparity and lower body weight (Boyd et al., 1998).

The potential role of commonly prescribed medications in the development of breast cancer is an emerging and growing area of breast cancer etiologic research (Aronson and Woolcott, 2001). It is well known that both endogenous (produced in the body) and exogenous (hormonal medications) female hormones affect the growth of mammary tissue and therefore play an important role in the development of breast cancer. Epidemiologic studies designed to examine risk associated with the use of oral contraceptive (OC) medications and hormonal therapy (HT), have yielded inconclusive results (ESHRE Capri Workshop Group, 2004). Changing patterns of use and formulations over time, different ages of study subjects as well as varying representations of exposure make it difficult to compare results between studies. However, evidence suggests that specific sub-groups of oral contraceptive users may be at increased risk including long-term users (>10 years) (VanHoften et al., 2000) and women with BRCA1 or BRCA2 mutations (Ursin et al., 1997). Similarly, sub-groups of HT users with elevated risk include long-term users (≥5 years) (Biglia et al., 2004), users of combined estrogen-progestin therapy (versus estrogen alone) (Collins et al., 2005) and women with a family history of breast cancer (Biglia et al., 2004). For both OC and HT users, risk decreases with increasing time since the medication was last used (ESHRE Capri Workshop Group, 2004; Collins et al., 2005).

In addition, there is laboratory and epidemiologic evidence (though inconsistent) to suggest that SSRIs (see complete literature review below), tricyclic antidepressants
(Sharpe et al., 2002), statins (Beck et al., 2003) and some anti-diabetic drugs (sulfonylureas and insulin) (Bowker et al., 2006) may be associated with increased risk of breast cancer. On the other hand, non-steroidal anti-inflammatory drugs (NSAIDS) (Sharpe et al., 2000; Cotterchio et al., 2001) and other anti-diabetic drugs (metformin) (Evans et al., 2005) may be associated with decreased risk.

Despite a wealth of epidemiologic literature related to breast cancer risk factors, the etiology of breast cancer remains largely unexplained, as well-established risk factors (specifically including family history of breast cancer in a first-degree relative, later age at first birth, nulliparity and higher family income) have been estimated to account for less than 50% of all cases of breast cancer (Madigan et al., 1995). This highlights the need to investigate other potential risk factors.

Further, breast cancer is a heterogeneous disease where age, menopausal status and stage at diagnosis as well as histopathologic subtypes may reflect different etiologies and risk factor profiles (Brinton et al., 2002; Tryggvadóttir et al., 2002). Specifically, there is evidence to suggest that the strength and direction of association between breast cancer and specific reproductive, anthropometric, lifestyle and dietary factors differ according to age at diagnosis (Tryggvadóttir et al., 2002; Russo, 2002; Yankaskas, 2005-2006; Key et al., 2001). Similarly, certain risk factors (including oral contraceptive use, reproductive factors and smoking) may be associated with different magnitudes of risk for pre- and post-menopausal breast cancer (VanHoften et al., 2000; Tryggvadóttir et al., 2002; Hulka and Moorman, 2001; Yankaskas, 2005-2006; Biglia et al., 2004).
2.3 SSRIs and breast cancer

2.3.1 Experimental Research

Experimental evidence related specifically to the potential carcinogenic effects of SSRIs is limited to a few *in vivo* and *in vitro* studies that mainly looked at the effects of fluoxetine on tumours. In a widely quoted 1992 study, Brandes et al. (1992) reported an accelerated growth rate of induced mammary tumours in rats exposed to clinically relevant concentrations of fluoxetine, decreased latency to the appearance of tumours by 30 to 40% and increased tumour frequency up to 2.5 – fold, as compared to controls who were administered saline only (no fluoxetine). In addition, Brandes et al. (1992) demonstrated that fluoxetine *in vitro* inhibited the binding of histamine to cytochrome P450 enzymes in porcine (swine) liver nuclei, inhibited normal lymphocyte proliferation in fresh spleen cell cultures and stimulated DNA synthesis in fibrosarcoma cell cultures (Brandes et al., 1992; LaBella and Brandes, 1996).

Wright et al. (1994) reported that fluoxetine inhibited DNA fragmentation and suppressed apoptosis (the normal mitochondrial-mediated physiological process of cell death) of seven different cancer cell lines (including human breast cancer cells) with subsequent enhanced proliferation of abnormal cells. Results from an *in vitro* study by Nahon et al. (2005) investigating the role of fluoxetine on apoptosis suggested that fluoxetine interacted with a key regulator of apoptosis, the mitochondrial component VDAC (voltage-dependent anion channel), preventing apoptotic cell death. Further, fluoxetine, administered to rat pups subjected to maternal separation, was shown to enhance cell proliferation and to suppress apoptosis in dentate gyrus brain tissue (Lee et al., 2001).
However, there are also *in vivo* and *in vitro* studies that have reported potential inhibitory effects of fluoxetine (SSRIs) on cell growth and proliferation (Volpe et al., 2003; Nordenberg et al., 1999). Contrary to the experimental studies by Brandes et al., (1992), Bendele et al. (1992) (from the research laboratories of Eli Lilly pharmaceutical company) reported no evidence to suggest that fluoxetine influenced spontaneously occurring breast tumour growth in rats. Also, using *in vitro* assays of cell proliferation with human and murine tumour cell lines, Volpe et al. (2003) reported significant suppression of DNA synthesis and a lack of direct stimulation of tumour cell proliferation with fluoxetine. In addition, there is experimental animal and laboratory research that demonstrates inhibitory cellular proliferative effects of fluoxetine on the growth of colon (Tutton and Barkla, 1982) and prostate tumour cell lines (Abdul el al., 1995).

This lack of consistency between experimental studies examining the effects of SSRIs on cell proliferative activity may reflect different study methodologies that tested (a) different cell types (human versus murine; neoplastic versus normal cell lines; different cancer cell types), (b) a wide range of different dosages (equivalent human doses versus higher doses) and (c) varying formulations, durations and routes of SSRI exposure (intravenous, oral or topical) (Volpe et al., 2003; Miller, 1995).

### 2.3.2 Epidemiologic Research

Epidemiologic research specifically looking at the relationship between SSRIs and breast cancer risk is limited to six case-control studies that collected self-reports of SSRI use from questionnaires, and two case-control studies and three retrospective cohort studies that captured SSRI exposure information from prescription or physician databases.
**Case-control studies:** Kelly et al. (1999), in a hospital-based case-control study (hospitals from North-Eastern United States) using questionnaires to ascertain SSRI exposure, reported no statistically significant association between either regular use of SSRIs (defined as use at least four days/week for at least four weeks) or duration of regular use (represented as the lifetime sum of periods of regular use) and breast cancer risk. However, an elevated risk of “borderline” statistical significance was observed associated with recent, regular SSRI use within the year prior to diagnosis of breast cancer (OR=1.8, 95% CI: 1.0-3.3), controlling for many potential confounders. Results were generally similar using either of two hospital control groups (cancer and non-cancer groups), suggesting that selection of controls was independent of the exposure of interest and that recall bias was not a limitation. However, sample size was small (22 cases and 23 controls) and the use of SSRIs of more than three years could not be assessed as most subjects were enrolled before SSRIs were approved and marketed (Kelly et al., 1999).

In a population-based case-control study (Ontario, Canada) by Cotterchio et al. (2000) results based on a self-administered questionnaire suggested that the use of paroxetine (defined as use for more than two weeks started at least six months prior to diagnosis) may be associated with a “substantial increase in breast cancer risk” (multivariate OR=7.2, 95% CI: 0.9-58.3). However, the number of paroxetine users was very small (nine cases and one control). Steingart et al. (2003), in a population-based case-control study (Ontario, Canada) which used a previously validated self-administered questionnaire and methods to limit recall bias, reported marginally statistically significant age-adjusted odds ratios for the use of sertraline (OR=1.58, 95% CI: 1.03-2.41) and paroxetine (OR=1.55, 95% CI: 1.00-2.40): use was defined as taken daily for at least two months and started at least 12 months prior to diagnosis. Multivariate analysis adjusting
for several confounders yielded similar but not statistically significant odds ratios. No association was observed for timing (that is, number of years since first or last use of any antidepressant) or duration of use for all antidepressants combined (TCAs, SSRIs and MAOIs) (Steingart et al., 2003).

Moorman et al. (2003) analyzed data from in-person interviews of women enrolled in the Carolina Breast Cancer and Carcinoma in situ Study and reported that the use of SSRIs for 36 months or longer increased the risk of invasive breast cancer (age and race-adjusted OR = 2.2; 95% CI=0.8-6.3); however, there were only 13 cases and 5 controls with this duration of exposure. Coogan et al. (2005) in an update of the Kelly et al. (1999) Ontario hospital-based case-control study using data from in-hospital interviews collected between 1988-2002, reported no association between breast cancer risk and (a) the regular use of all SSRIs combined (regular use defined as use at least four days/week for at least three continuous months), (b) regular use of three specific individual SSRIs (paroxetine, fluoxetine and sertraline) and (c) four or more years of SSRI use (estimated as the sum of all periods of regular use).

González-Pérez and Rodríguez (2005), in a nested case-control data linkage study, collected prescription and cancer diagnostic information from the General Practice Research Database, a computerized information system that contains data entered by general practitioners in the United Kingdom for more than two million patients. Current long-term use of SSRIs (as defined by more than three years of use that continued up to the index date or ended in the year prior to the index date) was not associated with an increased risk of breast cancer controlling for a limited set of potential confounders (OR = 0.56, 95% CI: 0.27-1.18): results are based on 7 cases and 80 controls in this exposure category. Since this database records out-patient prescription information based only on
general practitioner’s records of each drug prescribed at each appointment and therefore does not reflect prescriptions that were actually filled at a pharmacy, there may be greater potential for nondifferential misclassification of exposure as compared to studies that collect drug information from a pharmacy-based prescription database.

Chien et al. (2006) conducted a population-based case-control study of Washington DC women ages 65-79. The study population consisted of 975 breast cancer cases diagnosed between 1997-1999 and 1007 controls, frequency matched to the cases on age, year, and county of residence. Information on antidepressant use was ascertained using structured in-person interviews that included specific tools to enhance subject recall such as pill name lists and life events calendars. No elevated risk of breast cancer was reported for women who had ‘ever’ used SSRIs at any time prior to their diagnosis/reference date (‘ever’ use defined as use for three or more months). A sub-group analysis of women with a first degree family history of breast cancer suggested that for SSRI users (defined as use for three or more months) there was a reduced risk of breast cancer as compared to nonusers of SSRIs (OR = 0.4, 95% CI: 0.2-1.0). Alternatively, for SSRI users without a first degree family history, there was an increased risk reported compared to nonusers of SSRIs (OR = 1.5, 95% CI: 1.1-2.0). In addition, an elevated risk of progesterone receptor negative (PR-) and estrogen receptor positive/progesterone receptor-negative (ER+/PR-) breast cancers was reported for users of SSRIs (three or more months) as compared to nonusers (Chien et al., 2006).

A record linkage population-based case-control study by Fulton-Kehoe et al. (2006) that included 2,904 breast cancer cases diagnosed between 1990 and 2001 and 14,396 age-matched controls, all women enrolled in the Group Health Cooperative (GHC) of Washington state, ascertained antidepressant use from the Group Health
Cooperative (GHC) pharmacy database. Information on risk factors for breast cancer and other potential confounders was collected from questionnaires regularly administered as part of the GHC Breast Cancer Screening Program. No increased risk was reported for long-term use of SSRIs combined (as defined by 21 or more prescriptions) or for long-term use of individual SSRIs (as defined by 11 or more prescriptions). Contrary to Chien et al. (2006) no association was found between SSRI use with hormone receptor status sub-groups, and a first degree family history of breast cancer did not modify the association between SSRI use and breast cancer risk.

**Retrospective Cohort studies:** Dalton et al. (2000), in a Danish population-based cohort study, linked the Prescription Database of the County of North Jutland, Denmark with the Danish Cancer Registry to estimate the relative risk for all types of cancer after antidepressant use. Among the cohort of 13,015 users of SSRIs identified in the database between January 1989 and December 1995, no clear indication of an increased risk for any type of cancer was observed. However, the authors suggest caution when interpreting these results since the follow-up time for SSRI users was very short (mean of 1.6 years beginning one year after the date of the first known prescription) (Dalton et al., 2000).

In a retrospective cohort study by Wang et al. (2001) using the New Jersey Medicaid and the New Jersey Pharmaceutical Assistance to the Aged and Disabled (PAAD) program prescription databases for exposure information, no statistically significant association between breast cancer risk and fluoxetine use (defined as having filled at least one prescription) was observed controlling for a limited set of confounders available in the databases (adjusted hazard ratio=1.04, 95% CI: 0.87-1.25). Exposure
status was based on prescription records between 1989 and 1991 and subjects were not contacted to assess use after 1991 or to collect data on other potentially important confounders.

Results from a retrospective cohort study by Haque et al. (2005) designed to specifically estimate breast cancer risk associated with paroxetine use among a cohort of antidepressant users (users of SSRIs and atypical antidepressants) identified in the Kaiser Permanente Southern California (KPSC) database between 1995 and 2000, reported no statistically significant increased risk of breast cancer among women who exclusively used paroxetine relative to those who used other antidepressants controlling for age only (OR = 1.16; 95% CI: 0.98-1.38). Use of paroxetine for two years or more was not associated with an increase in breast cancer risk compared to those who used the medication for less than two years.

Reviews: Bahl et al. (2003), in a review of the literature in the MEDLINE database published from 1970 to 2002, concluded that the evidence concerning breast cancer risk and the use of antidepressants in general was inconclusive, but further research evaluating the use of SSRIs “may be of importance.” Lawlor et al., (2003), conducted a systematic review of eight epidemiologic studies published up to 2002 and combined trial data from one pharmaceutical company (Eli Lilly Limited) to assess the association between antidepressant use and breast cancer. A formal meta-analysis was not performed due to differences between the studies in terms of design, exposure representations and outcomes. The authors concluded that overall the evidence did not support an association but suggested that the relatively short durations of exposure considered in these studies may not be sufficient to detect an effect of exposure on breast cancer development.
Coogan (2006), in a review of the epidemiologic literature on the association between antidepressant use (including SSRIs) and the risk of breast cancer published between 1981 and 2005 concluded that overall, the evidence did not support the hypothesis that antidepressant use increased the risk of developing breast cancer. However, the author recommended that due to a ‘dearth of data’ on long-term SSRI use, that breast cancer incidence be monitored among women using SSRIs for long periods of time.

**Summary of the epidemiologic research:** The 11 epidemiologic studies that have examined the relationship between SSRIs and breast cancer risk have produced conflicting and inconsistent results ranging from marginally statistically significant odds ratios (Kelly et al., 1999; Cotterchio et al., 2000; Steinghart et al., 2003; Moorman et al., 2003; Chien et al., 2006) to findings of no significance (Dalton et al., 2000; Wang et al., 2001; González-Pérez and Rodríguez, 2005; Haque et al., 2005; Coogan et al., 2005; Fulton-Kehoe et al. 2006). These inconsistencies may reflect differences in study design, variations in exposure measurement categories used to represent dosage and duration of SSRI use, varying referent group populations used for comparisons, and uncontrolled or residual confounding. Considering these important variations in methodology, it is difficult to compare results between studies.

In addition, most of these studies were conducted in the mid to late 1990’s when SSRI use was less prevalent and therefore many of these studies were limited by a lack of ‘highly exposed’ subjects with prolonged use of SSRIs. Also, the potential effects of SSRI exposures on breast cancer development may be such that more time is needed due to latency issues in order to detect an association. In addition to a lack of detailed and precise characterization of risk in terms of long-term SSRI exposures during meaningful
exposure windows, limitations of previous studies included the potential for recall error and bias with self-reported SSRI use and limited information on established risk factors for breast cancer that may confound the risk estimates.

Two review articles (Bahl et al., 2003; Coogan, 2006) recommend further study related to long-term SSRI use and breast cancer risk. Considering that SSRIs have now been on the market for more than 18 years and are increasingly widely prescribed, further investigation of this potentially important biologically plausible relationship within the context of a study design that addresses the limitations of previous epidemiologic research, is justified to ascertain whether a relationship between SSRIs and breast cancer exists or not.

2.4 Underlying biologic mechanisms

Multiple complex biologic mechanisms based on animal, laboratory and epidemiologic research have been proposed to explain a potential tumour-promoter role of SSRIs in breast cancer development. Evidence suggests that SSRIs enhance tumour growth by disrupting normal cell growth regulatory mechanisms during a relatively late phase in breast cancer development, either indirectly through an SSRI-mediated increase in prolactin (a hormone for which there is increasing evidence to suggest an association with breast cancer), or directly through SSRI tumour-promoting mechanisms.

2.4.1 SSRI prolactin hypothesis

There is considerable animal and in vitro laboratory evidence to provide a biologic basis for an association between elevated prolactin levels and breast cancer development in humans (Harvey, 2005; Vonderhaar, 1999; Emiliano and Fudge, 2004; Welsch and
Nagasawa, 1977; Kiss et al., 1987). Prolactin, as a tumour promoter, has been shown to stimulate proliferative activity in the mammary gland, to suppress apoptosis (normal process of cell self-destruction) and to upregulate BRCA1 (Harvey, 2005; Vonderhaar, 1999). All five SSRIs have been shown to be associated with increased prolactin levels in humans (Emiliano and Fudge, 2004). SSRIs, (secondary to increased serotonin levels at neuronal synapses) stimulate the release of prolactin releasing factors (PRFs) which directly stimulate lactotrophs in the pituitary gland to release prolactin (Emiliano and Fudge, 2004). Petit et al. (2003) investigated the rates of hyperprolactinemia induced by prescription medications (including SSRIs) using adverse drug reaction reports documented in the French Pharmacovigilance Database between 1985 and 2000. A statistically significant increased risk of hyperprolactinemia was observed with all SSRIs except citalopram (Petit et al., 2003). Emiliano and Fudge (2004), in a review of the role of SSRIs in hyperprolactinemia, recommended further ‘controlled’ studies where pre- and post-SSRI-treatment prolactin levels are measured in well-characterized patient groups.

If a prolactin mechanism may explain a carcinogenic effect of SSRI use, then it is of interest to look at the epidemiologic literature related to elevated prolactin levels and the subsequent development of breast cancer. Three prospective case-control studies nested within the Nurses Health study cohort have examined this relationship. Hankinson et al. (1999) reported a statistically significant increased risk of breast cancer among postmenopausal women who had prolactin levels above 9.7 ng/ml (relative risk of 2.03, 95% CI: 1.24-3.31); mean time from blood collection to diagnosis/reference was 27.8 months (range from less than one month to 57 months). In another large nested case-control study, Tworoger et al. (2004) observed a modest increase in risk of postmenopausal breast cancer comparing the highest versus the lowest prolactin
concentration quartile (RR = 1.34, CI: 1.02-1.76): the association appeared to be strongest among cases diagnosed within two years of blood collection. Further, Tworoger et al. (2006) reported a modest positive association between prolactin levels and breast cancer risk predominantly among premenopausal women (RR = 1.5, 95% CI: 1.0-2.3), comparing the highest quartile versus the lowest quartile of prolactin levels. The association was stronger for cases diagnosed within approximately 3.9 years of blood collection compared to cases diagnosed more than 3.9 years after the blood draw, although differences were not statistically significant (Tworoger et al., 2006).

In summary, all SSRIs have the potential to elevate prolactin levels. Hyperprolactinemia has been shown to be associated with breast tumour growth and progression in experimental studies, and has been linked to increased breast cancer risk in epidemiologic studies.

2.4.2 Other SSRI tumour-promoting mechanisms

In addition to the SSRI prolactin hypothesis, there is experimental evidence to suggest that SSRIs may promote tumourigenesis directly by modifying breast cell proliferative activity during a late stage of breast cancer development. Several mechanisms have been proposed to explain this potential tumour promoter effect of SSRIs. One potential mechanism relates to the activity of the cytochrome P450 enzymes, an important family of microsomal enzymes that are found in all cells, especially the liver. These intracellular enzymes control the steady-state intracellular levels of lipid hormones that are responsible for regulating cell growth (Labella and Brandes, 2000). LaBella and Brandes (1996) postulated that fluoxetine, due to its structural similarity to an intracellular histamine receptor ligand, N,N-diethy-2-[4-
(phenylmethyl)phenoxy]ethanamine HCl (DPPE), a potent stimulator of tumour growth in vivo, interacted with intracellular growth-regulatory histamine receptor sites on cytochrome P450 monooxegenases resulting in increased cytochrome P450 enzyme catalytic activity. This increased cytochrome P450 activity altered the normal steady-state intracellular levels of lipid mediators responsible for regulating cell growth, thus leading to enhanced cell proliferation and tumour growth (LaBella and Brandes, 1996).

There is animal and laboratory evidence to suggest additional tumour promoter properties of SSRIs including: (a) stimulation of DNA replication and decreased DNA repair ability leading to unregulated DNA synthesis (Brandes et al., 1992), (b) suppression of the immune system, thereby diminishing the natural ability of the immune system to eliminate abnormal cells that have the potential to become neoplastic (Brandes et al., 1992) and (c) suppression of apoptosis (the normal process of cell death) resulting in enhanced survival of abnormal cells that have the potential to progress to cancer cells (Wright et al., 1994).

2.5 Breast cancer development

2.5.1 Overview of carcinogenesis and tumour promotion

Carcinogenesis is a multi-step process consisting of two main phases, initiation and promotion. Figure 1 (Alberts et al., 2002) is based on evidence from laboratory studies investigating the effects of cancer-causing chemicals on mouse skin but can be applied to other cancer sites. It illustrates potential cancer and non-cancer outcomes in relation to the timing of exposure to a tumour initiator (defined as a mutagenic substance which causes irreversible damage to a cell’s DNA creating an ‘initiated’ cell) and a tumour promoter (defined as a substance which is not mutagenic on its own but may
enhance proliferation of a previously ‘initiated’ or mutated cell) (Alberts et al., 2002). As Figure 1 indicates, in order for cancer to develop, exposure to a tumour promoter must be characterized by prolonged and regular use after a mutagenic event (initiation) has occurred (Alberts et al., 2002). Tumour promoters are postulated to enhance carcinogenesis by activating the expression of previously mutated growth-controlling genes or by suppression of growth-inhibitory influences such that the mutated cell is allowed to replicate, growing into a larger clone of cells with the potential for further chance genetic mutations and subsequent cancer development (Alberts et al., 2002).

**Figure 1: Initiation and promotion of cancer: schematic representation of experiments performed on mouse skin**

![Diagram](https://via.placeholder.com/150)

2.5.2 Conceptual model of exposure to SSRIs and the development of breast cancer

Figure 2 presents a simplified conceptual model for a relationship between SSRI exposure and breast cancer development. This model has been modified to include exposure to SSRIs, and was originally proposed by Love et al. (2002) to describe the process of molecular and physiologic events that lead to critical genetic abnormalities in a
group of malignant breast cancer cells. During the initiation phase, exposure to a carcinogen (initiator) may cause permanent and irreversible DNA damage to a breast cell creating a so-called “initiated cell” which may remain dormant until exposed to a promoting agent. It is during the promotion phase, a relatively late phase in breast cancer development, when SSRIs have been postulated to enhance proliferation of this ‘initiated’ or mutated cell.

Moolgavkar et al. (1980), using mathematical models, estimated the average time between the end of the promotion phase (when malignant cells are present) and the presence of clinically detectable breast cancer (the so-called “silent” interval) to be about five years, within a model where the data also fit well within a 2.5 to 7.8 year period of time. Laboratory research by Brandes et al. (1992) suggests that exposure to SSRIs affects a late stage of breast cancer development (promotion phase), therefore potentially shortening this “silent” interval: exposure to fluoxetine accelerated the growth and development of induced mammary tumours in rats, resulting in a decrease in the time to appearance of mammary tumours of 30-40% compared to saline controls. Based on the proposed promotion factor effects of SSRIs on breast cancer development, a 2-year “silent” interval has been added to the conceptual model by Love et al. (2002) to indicate a potential two year time period between the end of the promotion phase of breast cancer development (by which time malignant growth is present) and the presence of clinically detectable disease.

In summary, based on potential tumour-promoter effects of SSRIs, it is plausible to consider the risk of breast cancer development among women who used SSRIs for prolonged periods of time more than two years prior to diagnosis.
Figure 2: Conceptual Model of Breast Cancer Development

Modified from Love et al., 2002

2.6 Pharmacoepidemiology: investigating the long-term safety of drugs

Before a drug is approved for general use in the United States and Canada, premarketing studies of drug effects (both intended and unintended/adverse effects) include preclinical animal testing, and subsequent clinical trials that evaluate a drug’s overall efficacy and safety over short periods of time among relatively small, homogeneous study populations (Strom, 2005, p.9). Once the drug is approved and marketed, further monitoring and evaluation of its long-term safety in diverse populations is essential. Spontaneous reporting of adverse effects provides limited information related to potential safety concerns due to external influences and potential bias (Reynolds et al., 2005, p.80).

Post-marketing observational epidemiologic research can address safety issues related to long-term drug exposures and rare outcomes with long latency periods (such as cancer outcomes) (Reynolds et al., 2005, p.79). Pharmacoepidemiology is “the study of the use of and the effects of drugs in large numbers of people” (Strom, 2005, p.3). A
primary focus of this relatively new and rapidly growing field of epidemiology is the study of the distribution and determinants of adverse drug effects including the investigation of biologically plausible, uncommon, delayed drug effects related to long-term drug use.

Over the last 25 years, the development of computerized automated prescription databases containing detailed long-term prescription drug data has facilitated the study of relationships between prolonged drug use and cancer risk. Strom (2005, p.220) states that this prescription drug data, which is usually collected for administrative purposes and generated automatically by pharmacy payment submissions or claims, is considered to provide ‘some of the best data on drug exposure in pharmacoepidemiology’.

However, even though accurate drug exposure data and cancer outcome information can be ascertained relatively easily by electronically linking large prescription and cancer databases, these databases are often limited by a lack of detailed information on various potential confounders such as lifestyle factors, smoking and reproductive history, information that can usually only be obtained directly from study subjects. Therefore, various methodological (two-stage design) and statistical (propensity score matching) approaches have been developed to enable control for confounding in the analysis of pharmacoepidemiologic studies (Csizmadi et al., 2005, p.804).

The two-stage case-control design was first described by White (1982) and further developed by Cain and Breslow (1988) as a cost-efficient approach to address the absence of information related to potential confounders in studies using large administrative databases for exposure and outcome data. Two-stage sampling is most useful when exposure information is relatively easily obtained (for example from a prescription database) but covariate data is more expensive to collect (for example, requires in-person
or telephone interviews) (Rothman and Greenland, 1998, p.112). Stage one of a two-stage case-control design consists of a record linkage study with exposure and outcome data ascertained from the complete study population. Stage two involves the collection of detailed confounder information from a small sub-set or fraction of the stage one cases and controls. Stage two subjects are selected according to both disease and exposure status using a “balanced design” approach whereby approximately equal numbers of cases and controls are selected from each exposure category, thus maximizing statistical efficiency (yields more precise measurements of association) by over-sampling the less common outcome-exposure categories (Collet et al., 1998; Cain and Breslow, 1988; Schaubel et al., 1997). In the analysis, first unconditional logistic regression is used to estimate risk using the stage two exposure-outcome data alone, controlling for confounders. Then the stage two parameter estimate and its variance are ‘corrected’ using specific formulas that take into account the proportions of stage one subjects (the stage two sampling fractions) from each exposure/outcome category with completed interviews (Collet et al., 1998). In this way, the final risk estimates include the exposure-outcome data collected for the complete stage one study group, account for the selection bias that was intentionally introduced by the stage two sampling process and are adjusted for potential confounders based on the confounder information collected from the stage two cases and controls (Cain & Breslow, 1988).

2.7 Drug exposure measurement issues

Within the rapidly growing field of pharmacoepidemiology, there is no true “gold standard” for drug exposure measurement, as both self-reported drug use and prescription database information have limitations in terms of their accuracy and reliability as sources
of exposure data (West et al., 2005; West et al., 1997). Questionnaires, which have been the main source of drug exposure information for epidemiologic studies (Kelly et al., 1999; Cotterchio et al., 2000; Steinghart et al., 2003; Moorman et al., 2003; Chien et al., 2006; Coogan et al., 2005), are considered to be limited by under- or over-reporting and inaccuracies in duration and timing of medication use (West et al., 1995; West et al., 2005). Boudreau et al. (2004) conducted a validation study within the context of a case-control study of breast cancer by Chien et al. (2006) to evaluate the accuracy of self-reported antidepressant prescription information using the Group Health Cooperative (GHC) automated pharmacy data as the ‘gold standard’ comparison. Study findings indicated only moderate recall accuracy for self-reports of antidepressant medication use, with a sensitivity of 49% for cases and 44% for controls during an eight-year period of time prior to the reference date.

Over the last 25 years, computerized databases (so-called ‘automated databases’) have been increasingly used as sources of information for pharmacoepidemiologic studies (Strom, 2005, p.220). Compared to questionnaires, prescription databases are considered more accurate sources of drug information related to long-term exposures and the timing of drug use (Strom, 2005, p.220-221; West et al., 1997; West et al., 1995), and these detailed data related to patterns of use are important in order to define exposure within the context of a carcinogenic hypothesis, that is when latency must be considered (Shapiro, 1989). Prescription databases allow exposure information to be collected on a large sample population more quickly and are usually less costly compared to questionnaires (Strom, 2005, p.219-220). In addition, selection and recall bias are minimized. However, compliance with the dispensed medications must be assumed (West et al., 2005, p.738).
All provinces and territories in Canada have some type of administrative population-based drug database which records out-patient prescription data, though the population and medications included in the databases vary greatly (CIHI website, 2000; Miller, 1996). Only British Columbia (since 1992), Manitoba (since 1995) and Saskatchewan (since early 1970’s) include virtually all residents in their prescription database (Miller, 1996), whereas other provincial databases only include specific sub-groups of the population such as residents 65 years of age and older or persons receiving social assistance (Miller, 1996). The Saskatchewan prescription database, which has been used extensively in studies of drug exposure and health outcomes (Downey et al., 2005, p.305), is the only Canadian population-based record of prescription drug use that has collected data for virtually all residents since 1989, the year SSRIs first became available in Canada. In addition, the SK drug database can be linked electronically to other databases including the population registry and the cancer registry.

2.8 Rationale for further study of the relationship between SSRIs and breast cancer

In summary, there are an insufficient number of epidemiologic studies with adequate sample size and detailed exposure data related specifically to SSRI use to judge whether there is or is not a relationship between SSRIs and breast cancer risk. A sufficient amount of time has passed since widespread introduction of SSRIs, and use of SSRIs is prevalent enough to facilitate a meaningful population-based study. Further, research specifically related to SSRI exposure is warranted to assess risk in relation to long-term duration and dosage of use during critical time periods of exposure. Designing a study that produces precise and valid drug exposure measurement is a challenge. In this study, objective and detailed exposure estimates will be obtained from a prescription database.
Given projections of increased SSRI use for several medical and psychiatric indications, assessing this biologically plausible relationship with breast cancer risk is important at this time.

2.9 Context of Thesis Research

I am a co-investigator on a CIHR grant (PI- KJ Aronson) to investigate the relationship between SSRIs and breast cancer risk in a two-stage case-control design. Figure 3 illustrates the two-stage case-control selection process. The large circle represents the underlying cohort for all study subjects, consisting of Saskatchewan women ages 28-79, eligible for prescription drug coverage for 10 or more consecutive years with no previous cancer diagnosis within 10 years of selection for this study. Stage one consists of a record linkage study utilizing two Saskatchewan (SK) health services databases and the SK Cancer Agency cancer registry. Stage one subjects include women diagnosed with primary invasive breast cancer between January 1, 2003 and December 31, 2007 identified from the SK cancer registry and ten controls per case randomly selected from the SK person registry system (population registry). SSRI exposure for all stage one cases and controls is being ascertained via linkage with the Saskatchewan prescription database. For stage two, information on potentially important confounders is being collected by telephone interview from a sub-sample of the stage one study population (200 cases and 200 controls) in order to better estimate the independent effects of SSRIs on breast cancer risk.

Throughout the development of this larger study, it has been the intent of the study team (J Ashbury, K Aronson, P Beck and W King) that I would focus on a sub-set of the complete stage one data for my M.Sc. thesis. In this regard I have taken the lead in
the conceptual development of research questions, study design and statistical analysis for this phase of the larger project.

**Figure 3: Overview of subject accrual process for two-stage study**

![Diagram](image-url)
Chapter 3.0  Methods

3.1  Objectives

The three objectives of this thesis research focused on developing a conceptual and statistical framework that would facilitate and drive the larger two-stage study investigation.

Objective 1: To conceptualize SSRI exposures
   a. To create representations for timing of SSRI exposures
   b. To create representations for duration and dosage of individual and combined SSRI exposures

Objective 2: To determine breast cancer risk associated with dosage, duration and timing of SSRI use among Saskatchewan women, controlling for a limited set of potential confounders

Objective 3: To determine the sampling methods for selecting stage two cases and controls

3.2  Overview of Thesis Research Design

This thesis project consists of a record linkage nested case-control study that utilizes three Saskatchewan (SK) health services databases, the person registry system (PRS) (also referred to in this thesis as the population registry) and the SK prescription drug plan database (SPDP database) which are both maintained by SK Health, and the cancer registry which is maintained by the Saskatchewan Cancer Agency (SCA). The nested case-control design refers to a case-control study that is conducted within (i.e. nested within) a well-defined source population or cohort (Rothman and Greenland, 1998,
This design was chosen as an alternative to analyzing the entire cohort of all eligible persons from the source population (who met the age, prescription coverage and no previous cancer diagnosis criteria) due to the fact that a smaller dataset was more manageable due to fewer potential mismatches and other problems arising during the linkage process that required manual verification.

The SK health services databases are recognized as valuable resources for pharmacoepidemiological research (Downey et al., 2005, p.306; Tilson, 1985). The person registry system (PRS) captures demographic and health coverage data for more than 99% of Saskatchewan residents excluding people whose health care is fully funded federally (members of the Royal Canadian Mounted Police, members of the Canadian Forces and inmates of federal penitentiaries). Since provincial legislation in SK mandates reporting of all cancer cases in SK, the Saskatchewan Cancer Agency (SCA) cancer registry database includes all residents of Saskatchewan diagnosed with cancer (Downey et al., 2005, p.300). Approximately 95% of cancer notifications are from pathology reports or specialist referrals and about 5% are from autopsies or death certificates (Downey et al., 2005, p.300).

The third database, the Saskatchewan prescription drug plan database (SPDP database), records outpatient prescription information for all Saskatchewan residents with the exception of approximately 9% of the population (primarily registered Indians), who have their prescriptions paid for by another government agency (Downey et al., 2005, p.300). The SK drug plan has generally operated as an income-based support program, whereby consumers are entitled to prescription benefits based on their income (Downey et al., 2005, p.296). Prescription data for drugs listed in the Saskatchewan Formulary is included in the database for all eligible residents of SK regardless of whether or not
prescriptions were paid for by the SK prescription drug plan. At the time of dispensing a prescription, the pharmacist electronically transmits prescription information from the pharmacy to SK Health where each transaction is verified for identification, and claimant and drug coverage eligibility (Downey, 2005, p.297). The results of this check are transmitted immediately back to the pharmacy, providing visual verification of the medication data at the pharmacy level at the time of prescription data entry (Downey, 2005, p.297).

Breast cancer risk associated with dosage, duration and critical time periods of SSRI exposure was assessed in case-control data identified from women in the Saskatchewan population eligible for outpatient prescription drug benefits. Objective and detailed exposure estimates were obtained from the prescription database (SPDP). Information on a limited set of established breast cancer risk factors that may confound this relationship was ascertained from the person registry system (population registry) and the prescription database. In order to shorten the time frame for data collection for this study, this thesis project was based on three years of prevalent data (2003-2005) of the complete stage one data (2003-2007).

3.3 Study Population

3.3.1 Source population

The source population (underlying cohort) for this study included women in the SK population eligible for outpatient prescription drug benefits who, during the study’s case accrual period (2003 to 2005), were aged 28 to 79, had prescription coverage for 10 or more consecutive years prior to their index date (defined below), and no previous cancer diagnosis within the past 10 years (prior to index date). Within the context of this
nested case-control study design, cases and controls were selected for this thesis using incidence density sampling. As described by Kelsey et al. (1996, p.190), in this type of case-control study, incident cases are identified from the underlying cohort of study-eligible subjects on an ongoing basis as they occur. Controls are sampled over the same period of time from the same cohort as for cases. At the time of selection, controls do not have the disease but may become a case later over the course of the study.

Further, the risk-set sampling approach was used to select cases and controls for this study (Rothman and Greenland, 1998, p.98). Briefly, the source population (cohort) was conceptualized as consisting of six ‘risk sets’ based on dividing the three-year study period into six sampling periods of equal duration. From each of these risk sets of study-eligible individuals, all cases were identified and a random sample of controls was selected. The details of this sampling strategy in relation to this study are described below.

3.3.2 Summary of case-control sampling process

The following section describes the case-control selection process in more detail as it relates to this study (Figure 4). For a description of the exact linkage procedures, please see Appendix 2.

For the purposes of this study, the dosage, duration and timing of SSRI exposures were defined in terms of each subject’s index date. For cases, the date of breast cancer diagnosis was the index date. The index date for each control was defined by a randomly chosen date corresponding to each six-month case accrual period (explained in more detail below).
Cases: Primary invasive breast cancers diagnosed between January 1, 2003 and December 31, 2005 were identified by the SCA using pathology reports, physician records, and hospital and cancer clinic charts received by the cancer registry. In total, 1,935 women diagnosed between 2003 and 2005 with ICDO code = C50 (breast cancer) and behaviour = 3 (malignant) were identified. Further, the SCA ascertained case eligibility according to prior cancer diagnosis within the previous ten years reported in the cancer registry. Women with a cancer diagnosis (other than non-melanoma skin cancer and carcinoma in situ of the cervix) in the 10 years prior to the breast cancer diagnosis date were excluded in order to insure that the case group was not ‘contaminated’ with women with recurrent breast cancers or breast metastases, cancers that may represent a different etiology and therefore a different relationship with SSRIs as compared to new primary breast cancers. Also, it is possible that the relationship between SSRIs and breast cancer might be different for a woman diagnosed with a second cancer (other than breast cancer) in the previous 10 years as compared to a woman with no cancer in the past 10 years.

This SCA breast cancer file was forwarded to SK Health when case eligibility according to exact age (28-79 at index date) and prescription coverage criteria were ascertained by SK Health through linkage with the person registry system (PRS). Women with less than 10 years of continuous prescription coverage prior to index date were excluded. In this way, all subjects could potentially have at least 10 years of drug prescription data recorded in the SK prescription database (depending on whether or not prescriptions had been filled). Therefore, after all exclusions, the eligible case file consisted of 1,273 women (66% of total number of breast cancer cases diagnosed in 2003-2005) who met the eligibility criteria of the source population (index date from
2003-2005, age 28-79, at least 10 years of coverage and no cancer diagnosis in previous 10 years).

Cases from the eligible case file were then divided into six sub-sets based on their index date (January 1 to June 30, 2003, July 1 to December 31, 2003, January 1, 2004 to June 30, 2004, etc… July 1 to December 31, 2005). For example, the first sub-set of cases consisted of women diagnosed between January 1, 2003 and June 30, 2003 who were age 28-79 and had 10 or more years of prescription coverage prior to diagnosis and no prior cancer diagnosis in previous 10 years prior to index date.

Controls: During the initial sampling phase, a potential control pool was selected by SK Health personnel from the person registry system. This group of 262,697 women, which will be used to select controls for the complete stage one sample, included all SK women born from 1923-1979 (inclusive) with 10 years of continuous prescription coverage starting on or before December 31, 1997. For the control selection process, this potential control pool was divided into six sub-sets based on the sampling time period corresponding to each of the six case accrual periods (January 1 to June 30, 2003, July 1 to December 31, 2003, etc… July 1 to December 31, 2005).

For example, the sub-set of potential controls corresponding to cases with index dates from January 1, 2003 to June 30, 2003 was identified first and consisted of all subjects from the potential control pool age 28 by January 2003 and less than or equal to age 79 by June 2003 with prescription coverage for at least 10 years. Up to this point, it was not possible in the control selection process to apply the criteria of no previous cancer diagnosis in the 10 years prior to index date. Therefore, this sub-set of potential controls was linked with cancer registry data and those subjects with a prior cancer diagnosed between January 1993 and June 2003 were excluded.
In the next step, 10 controls per case were randomly selected from this sub-set of potential controls without prior cancer, frequency-matched in five-year groups to the actual age distribution of cases with index dates from January - June 2003. Sub-sets of potential controls were selected in a similar manner for each successive six-month case accrual period and 10 controls per case randomly selected from each sub-set. As mentioned above, the index date for each control was defined by a randomly chosen date within each six-month accrual period. Controls were selected from each sub-set without replacement within each six-month accrual period. However, selected controls were placed back into the potential control pool for control sampling for subsequent six-month accrual periods, meaning that a control in an earlier six-month accrual period could be selected again to be a control or become case in a future six-month accrual period. This ‘continuous eligibility to become a case’ is an important feature of the incidence density design approach (Rothman and Greenland, 1998, p.98).

The final control group consisted of 12,730 subjects, 10 controls per case. Even though increasing the control to case ratio above 4:1 does not result in a meaningful improvement in the risk estimate (OR) that can be detected, given 80% power and a fixed case sample size (Schlesselman, 1982), a 10:1 control to case ratio was necessary to ensure an adequate control pool for the stage two control selection where response rates for controls were anticipated to be lower than 50%.
### 3.4 Data files

Extraction of data for this thesis within the context of the larger two stage study was carried out by employees of the SCA and the Research Services (RS) unit of SK Health according to their guidelines. The RS unit was responsible for the linkage and preparation of the four research datasets (subject file, demographic file, drug file and cancer file).
In Saskatchewan, the Health Services Number (HSN) is a lifetime number that uniquely identifies each Saskatchewan resident eligible for health coverage and enables linkage over time between the three databases (PRS, SPDP and SCA cancer registry) (Downey et al., 2005, p.295). In preparing the data files for this study, when PRS or SPDP information was linked to the SCA cancer registry, data were matched based on HSN, sex and date of birth (+/- two years). The vast majority of records matched on those three variables. For example, in order to avoid mismatches, records for breast cancer cases that did not match on HSN or that matched on HSN but not on sex or date of birth (n= 5 for this study) were verified manually (which involved matching on name). Three questionable matches for cases that could not be resolved manually were rejected.

During the preparation of these four data files, personal identifiers were removed and replaced by a study-specific identification number (Study ID number) as soon as linkage was complete.

3.4.1 Subject File (14,003 records): This file included all eligible breast cancer cases (1,273) and non-cancer controls (12,730) for the period of January 1, 2003 to December 31, 2005. The 14,003 records in this file corresponded to 13,450 unique individuals because some subjects were selected more than once during the study period: 498 subjects were selected as controls in two or more six-month accrual periods and 40 cases were also selected as controls prior to their breast cancer diagnosis date. Individuals who were included in the study more than once were assigned a different study identification number each time they were selected to be in the study.

Variables included in the subject file were study identification (ID) number, case/control status, birth year, enrolment date (defined as the later of September 1, 1975:
date that the SK drug plan started, or date of SK health coverage initiation), index date (for cases equals the breast cancer diagnosis date and for controls is a randomly chosen date within each six-month accrual period) and age as of index date.

3.4.2 Demographic data file (364,658 records): This file contained annual demographic data for the majority of subjects compiled for each year from enrolment date (defined above) to index date. There were three variables recorded in this file for each year demographic data was available: marital status, residence status (urban/rural) and income support status.

**Marital status:** In the demographic data file, there were three categories used to represent marital status: single, married/common-law, and ‘other’ which included separated, divorced or widowed women.

**Residence status:** This variable was dichotomized into two categories: urban and rural based on the residence code reported in the population registry for each data year.

Urban and rural residence status was defined as follows:

- **Urban:** This designation referred to persons living in either of Saskatchewan’s two census metropolitan areas (CMAs) (i.e., Saskatoon or Regina) as defined by Statistics Canada. CMAs are formed by one or more adjacent municipalities centered on a large urban area. The census population count of the urban core must be at least 100,000 to form a census metropolitan area. To be included in the CMA, other adjacent municipalities must have a high degree of integration with the urban core, as measured by commuting flows derived from census “place of work” data.
• Rural: This residence designation referred to persons living in all areas outside of either of SK’s census metropolitan areas.

**Income Support Status:** This variable indicated whether or not the subject was receiving income security benefits through the Government of SK Department of Community Resources and Employment or the Government of Canada during the data year. Not all of these programs were in place over the entire study period and eligibility criteria may have changed over time.

There are 4 categories of coverage status:

• No special coverage

• Saskatchewan Assistance Plan (SAP) coverage: This program provides assistance to families and individuals who, for various reasons, including disability, illness, low income or unemployment, cannot meet basic living costs.

• Family-based income security benefit programs:
  
  o SK Employment Supplement is a monthly payment that supplements the income earned by lower income parents and assists with child care costs while parents work

  o SK Child Benefit Program is a monthly allowance for children in lower income families. It assists lower income families with the costs of raising children and helps families remain in the workforce rather than relying on social assistance plan (SAP) coverage to meet their children’s basic needs.

  o Family Income Program: supplements the financial resources of low-income families with dependent children to assist with costs of food,
shelter and other necessities. Eligibility is based on annual family income and assets.

- Senior-based programs (senior defined as 65 years of age or older):
  - Guaranteed Income Supplement (GIS): provides additional money, in addition to the Old Age security pension, to low-income seniors living in Canada.
  - SK Income Plan (SIP) provides supplementary income for senior citizens who have little or no income other than the federal Old Age security pension and Guaranteed Income Supplement (GIS).

3.4.3 Drug File (1,146,745 records): This file contained prescription data from the outpatient prescription drug database (SPDP) for medications of interest dispensed between the subjects’ index date and enrolment date (defined as September 1, 1975 or date that SK Health coverage was initiated, whichever occurred later). Variables included in the drug file included drug name, dispensing date, dosage formulation (capsule, tablet, liquid, gel, patch, etc), quantity dispensed (number of units of each drug dispensed) and strength of each ingredient in each drug formulation. The actual dosage prescribed (number of pills to be taken each day) and the indication for prescribing the medication is not included in the SPDP database.

Although the SK Drug Plan places no limit on the quantities of drugs that may be prescribed, for most medications (including SSRIs) the contract between SK Health and each SK pharmacy allows pharmacists to claim a dispensing fee for each 34-day supply dispensed. Therefore most medications are dispensed as a 34-day supply. However, pharmacists may dispense more than one 34-day supply on a single day if necessary (to
accommodate travel etc) and the pharmacist is reimbursed accordingly. Medications for which pharmacists can claim a dispensing fee for each two month supply dispensed include oral contraceptives and some estrogens (See SK Drug Plan Formulary, 2006-07, Appendix E).

The prescription file did not contain information on prescriptions dispensed for the one and a half year period between July 1, 1987 and December 31, 1988 when SK Health converted temporarily to a family-based record system (rather than individual). However, this would not affect SSRI data collection for this study since SSRIs were not available in Canada until 1989. Prior to January 1, 1996, capture of prescriptions for nursing home residents may be incomplete because many facilities were on a global drug budget. In addition, the drug file did not include over-the-counter drug purchases, medications dispensed in hospital or samples from physicians.

Prescription information for 170 medications (using liberal criteria, drugs with evidence to suggest a relationship with breast cancer development) was included in the drug file. These medications can be grouped into eight main categories: antidepressants, oral contraceptives, hormone therapy, phenothiazines, non-steroidal anti-inflammatory drugs, benzodiazepines, corticosteroids and anti-diabetic medications. For the purposes of this thesis, only SSRIs, oral contraceptives and hormone therapy medications were included in the analysis.

3.4.4 Cancer file (1,305 records): This file contained cancer diagnostic information for the 1,273 breast cancer cases included in this study. There were 31 women with more than one breast cancer diagnosis recorded in this file: 30 women had two breast cancer diagnoses and one woman had three breast cancer diagnoses recorded on the same day for
a total of 32 (30+2) additional breast cancer diagnoses recorded. These individuals with more than one breast cancer diagnosis represented women diagnosed on the same day with 2 (or 3) separate tumours in separate quadrants of the breast. Variables included in the cancer file were diagnosis date, ICDO code (location of breast tumour site code), behaviour (all were 3 = malignant), TNM (tumour, node, metastasis) staging value, specific TNM_T, TNM_N and TNM_M values, ICDO_M (morphology code; i.e. cell type: epithelial, squamous, ductal, lobular, papillary), method of diagnosis and specific location of metastasis. Hormone receptor status information was not available in the cancer registry.

3.5 Outcome Assessment: Case definition

Detailed diagnostic information on primary invasive breast cancer cases was included in the case file obtained from the cancer registry of the SCA. As described above, variables included diagnosis date, ICDO code (location of breast tumour site code), behaviour (for all cases behaviour = 3; malignant), ICDO_M (morphology code) and TNM stage information. For the purposes of this thesis, inclusion in the case group as invasive breast cancer was based on a behaviour code of 3 (malignant tumour growth at the primary site). In Saskatchewan, cancers are staged using all relevant information from many sources including pathology reports, radiology reports and relevant notes/reports from physicians, hospital and cancer clinic charts. The SCA registry staff (coders) use all of this relevant information to determine the stage of the cancer according to standard guidelines. When the staging is completed, the behaviour code is also assigned by registry staff using standard rules to maintain consistency.
3.6 Exposure Assessment: SSRI exposure definitions

In previous epidemiologic literature investigating the relationship between SSRI use and breast cancer, SSRI exposure definitions ranged from imprecise dichotomous variables (Yes/No) to more sophisticated measures representing more precise amounts of SSRI exposure. For a carcinogenic hypothesis, it is important to define exposure in terms of regular use for long periods of time (preferably years) (Shapiro, 1989), and to know the timing of use. Therefore, it was important to create meaningful exposure representations of long-term SSRI exposures that considered timing, duration and dosage of use. SSRI exposure information for cases and controls was extracted as described above from the SPDP database for the period between: (a) January 1, 1989, the year SSRIs first became available in Canada; or (b) the date the subject first became eligible to receive benefits from SPDP (whichever was later), and the index date. As mentioned previously, data extracted from the SPDP database for each SSRI outpatient prescription dispensed included the dispensing date, the name of the SSRI, the strength and dosage formulation, and the quantity (number) of pills dispensed. The SPDP database does not include information on the daily dosage regime prescribed (pills per dose and total number of doses per day) and the individual indications for the prescription.

3.6.1 Characterization of SSRI exposures

Timing of SSRI exposures: (See figure 5) For all SSRI exposure variables, exposures occurring within two years prior to the index date were not considered due to the “silent” interval consideration mentioned previously. In this way, SSRI exposures not likely to be causally associated with breast cancer were excluded from the analysis. To study the effects of timing of SSRI exposures, risk was investigated in relation to three exposure
time periods. Firstly, risk was examined associated with total SSRI use during the main exposure window more than two years prior to index date. This time interval was divided into two shorter *a priori* time periods, 2-7 and more than seven years prior to index date. These two intervals were chosen to represent two biologically plausible periods of time when SSRIs, as potential tumour promoters may affect breast tumour growth. Also, the latency period between SSRI exposure and breast cancer development may be such that higher exposures more than seven years prior to index date may be associated with increased risk as compared to exposures during the period of time 2-7 years prior to index date or vice versa, relationships that might be missed if risk is examined only in relation to SSRI use over the entire study period.

Figure 5: SSRI exposure windows of interest

![Diagram showing exposure windows and index date](image-url)
**Categorical representations of SSRI exposures:** Overall, for each analysis, SSRI use was characterized as a categorical variable rather than a continuous variable in order to facilitate an analysis of the contrast of most interest in this study, that is, nonusers versus heavy long-term users. In addition, most subjects in this study were nonusers and a dose-response relationship was not necessarily anticipated. Therefore, SSRI use was characterized by categorical variables representing nonusers, light/sporadic users and heavy/regular users.

**Combined SSRI use:** Two measures of SSRI exposure were created to facilitate examination of risk associated with duration and cumulative dosage of combined SSRI use during the three *a priori* specified exposure time windows described above (>2 years, 2-7 years, >7 years prior to index date).

(a) **Duration of exposure variables:** Duration of exposure was represented by total number of SSRI prescriptions, which in Saskatchewan are typically dispensed to accommodate a 34-day treatment period. Exposure categories were chosen for each exposure window to ensure a high duration of exposure category with an adequate number of subjects. For example: for the overall time window more than 2 years prior to index date and the shorter exposure period 2-7 years prior to index date, exposure was classified as: never, 1-35 prescriptions (light users) and 36 or more prescriptions dispensed (heavy users). This cut-point for the highest exposure category was chosen to correspond with three cumulative years of SSRI use. Due to small sample sizes for highly exposed subjects for the period of time more than 7 years prior to index date, exposure was categorized as: never, 1-23 and 24 or more prescriptions dispensed (corresponding to two years of cumulative use).
(b) Dosage variables: To create a variable that represents combined SSRI dosage, total dosage was estimated in terms of fluoxetine-equivalent milligram dosages for each individual SSRI prescription dispensed and total, cumulative dosage calculated as the sum of all SSRI prescriptions of this equivalent dose. Equivalent dosage calculations were based on the defined daily dose (DDD) methodology published by the World Health Organization (WHO) to serve as a tool for drug utilization research (WHO Collaborating Centre for Drug Statistics Methodology, 2007). The DDD is a unit of measurement that allows comparisons of therapeutic intensity between medications from the same therapeutic class (Lee and Bergman, 2005, p.409). The DDD fluoxetine-equivalent dosages for the five SSRIs included in this study are: 20 milligrams (mg) of fluoxetine = 20 mg of citalopram = 20 mg of paroxetine = 50 mg of sertraline = 100 mg of fluvoxamine (WHO Collaborating Centre for Drug Statistics Methodology, 2007). Even though it is debatable as to whether or not fluoxetine-equivalent dosages are representative of the potential carcinogenic potency of SSRIs, using this measure to represent combined exposures has clinical relevance in that if an effect is observed, cumulative dosages can be interpreted within the context of true therapeutic dosages.

Dosage cut-offs for categories representing total dosage were based on examining SSRI exposure among the exposed control group. To form a relatively extreme uppermost dosage exposure category, the highest exposure category for each exposure window was chosen based on the total SSRI cumulative dosage cut-point of 10% of exposed controls. Total cumulative dosage of SSRI use was classified in the following categories: no exposure, light users (range of exposure from no exposure to cut-point defined above) and heavy users (total dosage greater than 10% cut-point). Risk was
examined in relation to total dosage of exposure during each of the three exposure windows (>2 years, 2-7 years, >7 years prior to index date).

**Individual SSRI use:** As outlined above, initially the analysis assessed risk associated with exposure to all SSRIs combined. However, since the effect of a more potent SSRI could be diluted by combining all SSRI exposures, it is important to assess risk associated with the use of individual SSRIs. In addition, even though SSRIs share the same therapeutic mechanism of action, namely serotonin reuptake inhibition, they each have different chemical structures and different pharmacologic actions (Stahl, 1998; Westenberg and Sandner, 2006). Also there is some evidence to suggest that individual SSRIs are associated with different adverse effects related to breast health, such as ductal carcinoma in situ (Markopoulos et al., 2005), fibrocystic breast disease (McKenzie and Risch, 1995; Hall, 1994; Amsterdam et al., 1997) and mammographic changes (‘bilateral hyperdensities’) (Marcus, 2001). Finally, the epidemiologic literature suggests that the use of paroxetine and sertraline may be associated with marginally increased breast cancer risk (Cotterchio et al., 2000; Steingart et al., 2003).

For each individual SSRI, duration of use was estimated using the number of prescriptions filled. Categories that included a highly exposed group with an adequate number of subjects were created to represent ‘exclusive’ and ‘any’ use of each individual SSRI during the main exposure window >2 years prior to index date. ‘Any’ users referred to women who had used the individual SSRI of interest alone or in combination with one or more other SSRIs during the main study period (more than two years prior to index date). ‘Exclusive’ users referred to a sub-set of ‘any’ users who had used only one specific type of SSRI during the study period.
**Paroxetine, sertraline and fluoxetine exposures combined:** Paroxetine is the most potent blocker of serotonin (5-HT) uptake at the presynaptic junctions in the brain (Turner et al., 2007). Sertraline and fluoxetine, though not as potent as paroxetine, also show a high degree of inhibition of serotonin uptake relative to fluvoxamine and citalopram. Given that this inhibition of serotonin uptake results in elevated serotonin levels at the neuronal synapse, and that serotonin has been shown to stimulate the release of prolactin releasing factors (PRFs) leading to increased prolactin secretion by the pituitary gland, it is biologically plausible that the use of these three relatively high inhibitors of serotonin reuptake (paroxetine, sertraline and/or fluoxetine) may be associated with an increased risk of breast cancer development. Therefore, a sub-analysis was done to test this hypothesis. Exposure variables were created based on the total number of prescriptions dispensed more than two years prior to index date for any of these three SSRIs.

### 3.7 Characterization of potential confounders

Figure 6 schematically represents the role of potential confounders on the association between SSRI use and breast cancer. Potential confounders in this relationship include protective effects and risk factors with established associations with breast cancer. Potential confounders *not* included in this thesis project, due to the fact that information was not available in the three databases, are highlighted in grey. For this study, information on a limited set of potential confounders (age, marital status, residence and income support status) was captured from the population registry. Prescription data on oral contraceptives and hormone therapy was collected from the SPDP prescription database.
Figure 6: Potential confounders for SSRI use and breast cancer risk

Note: Potential confounders that are highlighted in grey were not included in this study.

3.7.1 Confounders included in this study

Age: Age was calculated as the difference between the subject’s date of birth and index date. All multivariate analyses were adjusted for age.

Marital status: Marital status was considered as a potential confounder as it is associated with other established risk factors for breast cancer including reproductive factors (parity, age at first birth), health care utilization (mammography) and indicators of SES (income
level) (Roisman et al., 2004; Cabeza et al., 2007). For this analysis, there were three categories representing marital status recorded in the demographic data file: married, single and other (included widowed, separated or divorced). For the majority of subjects, marital status was assigned a category based on marital status six years prior to index date. For the 10 subjects with less than 10 years of demographic data, marital status was assigned based on data available in the demographic data file such that marital status was assigned based on status recorded five years before index date (if six years prior was not available), four years prior to index date (if five years prior was not available), three years before index date (if four years prior to index date was not available), two years prior to index date etc. etc. Three subjects only had marital status at index date and were categorized as ‘missing’.

**Socioeconomic status (SES):** The nature of the associations between breast cancer risk, SSRI use and socioeconomic status (SES) is unclear. Higher SES has been associated with increased breast cancer risk (Hulka and Moorman, 2001). There is some evidence based on U.S. study populations to suggest that individuals of higher SES may be more likely to be prescribed antidepressants. Sirey et al. (1999) reported that minority patients (Blacks, Hispanics) with depression were less likely to receive an antidepressant prescription as compared to whites with similar depressive symptoms. Also, Sleath & Shih (2003), using the U.S.1998 National Ambulatory Medical Care Survey (NAMCS), reported that patients with depression whose visits were covered by private insurance were almost twice as likely to receive an antidepressant prescription as compared to self-paying patients. In addition, younger patients, non-Hispanic whites and those with
private drug coverage were reported to be more likely to be prescribed an SSRI as compared to a non-SSRI antidepressant (Sleath & Shih, 2003; Olfson et al., 1998).

For this thesis project, SES was characterized using two indicators recorded in the SK person registry database, income support status and residence status (urban/rural status). Individuals of lower SES are more likely to have received government-funded income support. Evidence suggests that rural residents tend to be older age, and less educated with lower incomes as compared to urban residents (Monroe, 1992; Higginbotham, 2001). In addition, most studies have reported higher age-adjusted cancer incidence and mortality rates for most cancer sites in urban as compared to rural populations postulated to be due to increased exposure to environmental pollutants and higher prevalence of lifestyle factors associated with increased cancer risk such as smoking (Sheen, 1999; Monroe, 1992). However, rural residents relative to urban residents tend to present at initial diagnosis with a more advanced stage of cancer which is thought to be related to disparities in access to and utilization of health care services (Monroe, 1992; Higginbotham, 2001).

(a) Income support status: For the purposes of this thesis, the three main income support programs (Social Assistance Plan, the family-based income security benefit programs and the senior-based programs) were all considered to indicate some degree of lower SES and were therefore grouped together in the process of creating a variable to represent SES. It is important to note that not all of these programs were in place over the entire study period. Also, eligibility criteria for these benefits have changed over time.

Recognizing these potential limitations, using all of the yearly demographic data available, subjects were grouped into three categories based on income support received in the time period two or more years prior to index date; those who did not receive
income support benefits, those who received some coverage (defined as receiving any or all of the three types of income support for less than 50% of the years for which data was available) and thirdly, those who received a high level of income support (as defined by receiving any or all of the three types of income support benefits for more than 50% of the years for which data was available two or more years prior to index date).

(b) Residence status: Using the yearly residence information available in the population registry data file for the period of time two or more years prior to index date, subjects’ residence location was classified as urban if they resided in an urban location (population > 100,000) for more than 50% of the years for which data was available and rural if they lived in an urban location (population < 100,000) for less than 50% of the time. (No subjects had lived in rural and urban locations for an equal number of years.)

Therefore, as described above, SES was inferred on the basis of two indicators (income support status and residence status). Information on other potential confounding variables related to SES, such as income and education level were not available in the SK databases used for this thesis project.

Oral contraceptive (OC) and hormone therapy (HT) use: The relationship between the use of exogenous hormones and breast cancer risk has been studied extensively. Epidemiologic studies investigating breast cancer risk associated with the use of oral contraceptives (OCs) and hormone therapy (HT), have reported inconclusive results suggestive of increased risk with longer use (Dumeaux et al., 2004; VanHoften et al., 2000; Lee et al., 2005; McPherson and Mant, 2005). Based on this wealth of epidemiologic literature and the highly prevalent use of these medications, OC and HT use were chosen to be potential confounding variables for this project.
Medication history for OC and HT use was compiled for the period between (a) September 1, 1975 (year that SPDP database was started) or (b) the date of prescription coverage initiation, (whichever occurred later), and the subject’s index date. Consistent with SSRI exposure variables, use of OCs or HT within two years of index date was excluded from the analysis as exposures in the more distant past are more likely to be associated with cancer etiology.

(a) Oral contraceptives: Data pertaining to the type of formulation (oral or injectable), strength and quantity of each ingredient was ascertained from the prescription database. A dichotomous variable was created to represent any OC use with heavy users defined as having filled 12 or more prescriptions and light users as having filled 0-11 prescriptions during the period of time at least two years prior to index date.

(b) Hormone Therapy: Prescription data related to formulation type (oral, transdermal patch, vaginal ring or cream), strength of each ingredient and quantity dispensed up to index date was ascertained from the prescription database. Prescriptions for vaginal creams and rings were excluded since vaginal estrogen has been associated with poor compliance and there is evidence to suggest primarily local absorption rather than systemic effects of these HT preparations (Ayton et al., 1996).

For the remaining oral and transdermal HT formulations, a dichotomous variable was created to characterize OC use with heavy users defined as having filled 12 or more prescriptions and light users as having filled 0-11 prescriptions at any time at least two years prior to index date.
3.8 Sample size and power calculations

Based on the assumption that 3% of controls would have the highest SSRI exposure (defined \textit{a priori} as $\geq 36$ prescriptions) and 87% no exposure, sample size calculations included in the thesis proposal indicated that with 915 total cases (and approximately 824 cases in this contrast) and 10 controls per case, the study would have approximately 80% power to detect an odds ratio of 1.64 with two-tailed significance of 0.05.

3.9 Statistical Analysis

This study has two main objectives: to conceptualize SSRI exposures and to determine the association between dosage, duration and timing of SSRI use, controlling for a limited set of confounders. To achieve these objectives, the four data files described previously were linked by subject study identification number in order to combine drug data with demographic and cancer information.

3.9.1 Descriptive analysis

Prevalence figures over time for SSRI use combined and for individual SSRIs were calculated. Descriptive statistics were used to estimate the frequency distribution of a limited set of potential confounders (including demographic characteristics; age, marital status, urban/rural status and income support status, and the use of OCs and HT) in the case and control groups separately, and also in SSRI users and nonusers in the control group only.
3.9.2 Conceptualization of SSRI exposures

To achieve objective one, SSRI exposure variables were created to represent duration and total dosage of combined SSRI use and duration of use of individual SSRIs in relation to total overall use (more than two years prior to index date) and exposure during two specific shorter a priori exposure windows (2-7 years and more than 7 years prior to index date). The frequency distribution of individual and combined SSRI use among the case and control groups during each of the three exposure windows was also assessed.

3.9.3 Multivariate analysis

Assessment of confounders: As described previously, for this thesis project, information was available in the databases for the following potential confounding variables which have been shown in the literature to be related to breast cancer risk: age, marital status, income support status, residence status, OC and HT use. For this thesis project, two strategies for confounder selection were considered; a conventional regression method and the change-in-estimate method (Greenland, 1989).

For the main analysis, a regression method (backward stepwise elimination procedure) was used to create a parsimonious model of covariates that were predictive of breast cancer using a liberal p-value of 0.15. Since this confounder selection method depends only on the covariate having an effect on the outcome of interest and ignores covariate-exposure associations, a liberal significance level was used to avoid the exclusion of potential covariates that may have an effect on the exposure-outcome relationship but not be strongly related to breast cancer risk (Budtz-Jorgensen et al., 2006; Greenland, 1989). As described by Greenland (1989), conventional stepwise regression
starts with the fully adjusted model with age and exposure as forced-in variables (not subject to deletion) and the remaining potential confounders as variables subject to deletion. Specifically, for the stepwise backward elimination method, one by one, in a stepwise manner, the variable with the highest p-value is removed from the model until the point when the covariate with the smallest effect on the outcome of interest is associated with a p-value less than 0.15.

To provide confidence in the conventional regression method used for this study, a separate analysis was done to identify potential confounders using the change-in-estimate method as described by Greenland and Rothman (1998, p.257). This more efficient method of confounder selection is often recommended over methods based on significance testing (p-values) since covariates are chosen considering their effect on the exposure-outcome relationship of interest- which is arguably a better way of defining a confounder (Budtz-Jorgensen et al., 2006; Greenland, 1989). Starting with the fully adjusted model (containing all potential confounders), variables were evaluated in their roles in the SSRI/breast cancer relationship based on the change in the OR associated with the exposure contrast of most interest (nonusers versus those with 36 or more prescriptions) following omission from the model. At each step, the covariate showing the smallest change in OR when omitted from the model was removed from the model. After the removal of each covariate, the total percent change in the risk estimate (OR) between the new model (without the covariate) and the fully adjusted (original) model was calculated. Covariates were no longer removed from the model at the point when the total change in the OR accrued from the start of the process (with all confounders controlled) exceeded 10%.
Since confounder selection for the change-in-estimate method is based on the effect of each covariate on a specific exposure-outcome relationship, this method is less suitable when there are several exposure variables of interest. Each exposure representation would require a separate change-in-estimate confounder selection analysis with the possibility that each multivariate model for each SSRI exposure would include a different set of potential confounders making the interpretation of results cumbersome and problematic. Therefore, since there are several exposure representations of interest for this thesis project, a conventional stepwise regression method was used to identify potential confounders. The sub-set of covariates that were retained in the model using this approach was included as potential confounders in each multivariate analysis for each exposure representation of interest for this thesis.

**Multivariate risk analysis:** To examine the relationship between duration, dosage and timing of combined and individual SSRI use and breast cancer risk (objective 2), unconditional logistic regression was used to estimate the odds ratios and associated 95% confidence intervals adjusted for a limited set of potential confounders. The odds ratio was used as an estimate of relative risk.

**Interaction between age at diagnosis and SSRI exposure:** Evidence suggests that menopausal status at diagnosis may reflect different etiologies and risk factor profiles. Therefore, further analysis was undertaken to assess the association between SSRI use and breast cancer risk using age at index date as a proxy for menopausal status with pre-menopausal defined as 55 years of age or younger, and post-menopausal defined as over age 55 (VanHoften et al., 2000; Tryggvadóttir et al., 2002). To evaluate whether
menopausal status at index date modified the association between SSRI use and breast cancer risk, the significance of an interaction term (age at index date*SSRI exposure) was tested.

3.9.4 Analysis for stage two subject sampling

The third objective relates to stage two of the larger study, where a sub-set of stage one cases and controls is being selected for the purpose of assessing confounders, and detailed confounder information is being collected by telephone interview. Stage two subject sampling for the first set of telephone interviews was based on two of the three years of data (index date: 2004-2005) prepared for this thesis project. For this thesis project, a list of cases and controls to be invited for the first set of stage two interviews was compiled based on an analysis of the case and control SSRI exposure distribution for the 2004-2005 data.
Chapter 4.0  Results

4.1  Descriptive Analysis

4.1.1 Descriptive analysis of combined and individual SSRI use among controls only

The following results illustrate prescribing trends for all SSRIs combined and for each individual SSRI in the control group (as representative of the underlying study population) for the years 1989 to 2004.

Prevalence of SSRI use by year among controls

Figure 7 shows the yearly prevalence of SSRI use among controls from 1989 to 2004. For this analysis, SSRI users were defined as having filled one or more prescriptions for any SSRI in a given year. As predicted, the prevalence of SSRI use increased gradually from 0.2% of controls in 1989 to 7.1% in 2004. (See table 3.1 in Appendix 3 for exact data.)

Figure 7: Prevalence of SSRI use by year among controls
Prevalence of individual SSRI use from 1989-2004 among SSRI users (controls only)

Figure 8 illustrates the proportion of controls using each individual SSRI among all SSRI users (in the control group) from 1989 to 2004. SSRI users were defined as those who filled at least one prescription for any SSRI in a given year. Figure 8 illustrates the trends in SSRI prescribing practices between 1989 (when fluoxetine first became available in Canada) and 2004. Among SSRI users, the percentage of controls using each individual SSRI varied over the study period with a gradual rise in use for each SSRI after introduction into the pharmaceutical market and a gradual decline in use when a new SSRI became available. In the control population of SSRI users from 1989 to 1998, prevalence figures indicate that fluoxetine was the most commonly filled SSRI prescription overall. From 1999 to 2004, paroxetine surpassed fluoxetine as the most commonly prescribed SSRI medication. (See table 3.2 in Appendix 3 for exact data for each year from 1989-2004).

In Figure 8, total percentages for each year may exceed 100% to account for those controls who used more than one type of SSRI in a given year. For example, in 1999, of all SSRI users in the control group, 32.1% had used fluoxetine, 12.4% fluvoxamine, 24.5% sertraline, 37.2% paroxetine and 1.1% had used citalopram (total of 107.3%). Therefore, 7.3% of controls had filled prescriptions for more than one type of SSRI in that year.
Further analysis confirmed that controls had frequently filled prescriptions for more than one SSRI over the study period (see Table 1). Of the 1,857 controls that had filled one or more SSRI prescriptions over the study period, 1,291 individuals (69.5%) had filled a prescription for one type of SSRI, 435 (23.4%) for two SSRIs, 102 (5.5%) for three SSRIs, 26 (1.4%) for four SSRIs and 3 (0.2%) controls had filled prescriptions for all five SSRI medications over the study period.

Table 1: Use of one or more type of SSRI among controls

<table>
<thead>
<tr>
<th>SSRI use</th>
<th>Number of controls N=1857</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 type of SSRI</td>
<td>1291 (69.5)</td>
</tr>
<tr>
<td>2 types</td>
<td>435 (23.4)</td>
</tr>
<tr>
<td>3 types</td>
<td>102 (5.5)</td>
</tr>
<tr>
<td>4 types</td>
<td>26 (1.4)</td>
</tr>
<tr>
<td>All 5 types of SSRIs</td>
<td>3 (0.2)</td>
</tr>
</tbody>
</table>
**Frequency distribution of SSRI use among controls**

Table 2 and Figure 9 show the SSRI exposure distribution among controls during the main exposure window of interest, that is, during the period of time more than two years prior to index date. Exposure was represented by total number of prescriptions dispensed during this time interval. Of the 12,730 controls, 85.4% were nonusers, 9% had filled 1-11 prescriptions, 2.2% had 12-23, 1.1% had 24-35 and the remaining approximately 2.3% of controls had filled 36 or more prescriptions.

**Table 2: Frequency distribution of SSRI use among controls**

<table>
<thead>
<tr>
<th>Total number of SSRI prescriptions (more than 2 years prior to index date)</th>
<th>Frequency</th>
<th>% of all controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10873</td>
<td>85.4</td>
</tr>
<tr>
<td>1-11</td>
<td>1150</td>
<td>9.0</td>
</tr>
<tr>
<td>12-23</td>
<td>280</td>
<td>2.2</td>
</tr>
<tr>
<td>24-35</td>
<td>141</td>
<td>1.1</td>
</tr>
<tr>
<td>36-47</td>
<td>84</td>
<td>0.7</td>
</tr>
<tr>
<td>48-59</td>
<td>66</td>
<td>0.5</td>
</tr>
<tr>
<td>60-71</td>
<td>45</td>
<td>0.4</td>
</tr>
<tr>
<td>72-83</td>
<td>33</td>
<td>0.3</td>
</tr>
<tr>
<td>84-95</td>
<td>33</td>
<td>0.3</td>
</tr>
<tr>
<td>96-107</td>
<td>14</td>
<td>0.1</td>
</tr>
<tr>
<td>≥108</td>
<td>11</td>
<td>0.1</td>
</tr>
<tr>
<td>Totals</td>
<td>12730</td>
<td>Approx: 100%</td>
</tr>
</tbody>
</table>

Note: percentages may exceed 100% due to rounding
4.1.2 Descriptive analysis of cases and controls

Characteristics of cases and controls with respect to potential confounders

The study population included 1,273 women with primary invasive breast cancer and 12,730 controls, frequency matched to the cases by age. Of the 1,273 breast cancer cases, 415 (32.6%) were diagnosed in 2003, 429 (33.7%) in 2004 and 429 (33.7%) in 2005. All cases had a behaviour code of 3 (malignant tumour growth at the primary site) recorded in the cancer file. Cases represented a homogeneous group with respect to morphology code with approximately 75% of cases having infiltrating ductal carcinoma, 10% lobular carcinoma and 15% representing a mix of other morphologies. However, there was more variability in relation to stage at diagnosis with approximately 44% of cases presenting as stage one, 36% stage two, 12% stage three, 5% stage four and 3% unknown (i.e. stage could not be assessed due to lack of information at the time that the case file was compiled by the SCA).
Table 3 summarizes the frequency distribution for cases and controls, and unadjusted odds ratios (OR) and 95% confidence intervals (CI) for breast cancer with respect to selected potential confounders for which information was available in the SK databases. There was no significant difference in the mean age between cases and controls, with the majority over age 55 (about 65%), about 32% between ages 40 and 55, and about 3% under age 40. The mean length of enrolment in the SK Health-covered population eligible to receive outpatient prescription drug benefits, at about 28 years, was the same for cases and controls, with almost 99% eligible for prescription coverage for 15 or more consecutive years prior to index date. Case and control groups were also comparable in terms of income support status, residence status, and the proportion that were single. However, compared to controls, cases were slightly more likely to be married, and fewer cases were widowed, separated or divorced. Cases were slightly more likely to be heavy users of HT and/or OCs (as defined by 12 or more prescriptions more than two years prior to index date).

The unadjusted odds ratios for each potential confounder are also shown in Table 3. Marital status was associated with breast cancer risk with decreased risk observed for women in the widowed, separated or divorced category (OR = 0.80, CI = 0.69-0.93) as compared to women who were married. There is a suggestion of increased risk of breast cancer (not statistically significant) among women who were classified as heavy users of HT as defined by having filled 12 or more prescriptions (OR = 1.12, CI = 0.98-1.26) and/or OC (OR = 1.15, CI = 1.00-1.32) relative to those who were classified as light users of these medications. Unadjusted odds ratios indicated that income support status, residence status and age category were not associated with breast cancer risk. In
summary, marital status, OC use and potentially HT use were related to breast cancer risk in this study population.

Table 3: Characteristics of case and control subjects and unadjusted odds ratios (OR) and 95% confidence intervals (CI) for breast cancer associated with potential confounders

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases (N=1273)</th>
<th>Controls (N=12730)</th>
<th>Unadjusted OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>961 (75.5)</td>
<td>9196 (72.2)</td>
<td>1.00</td>
</tr>
<tr>
<td>Single</td>
<td>86 (6.8)</td>
<td>827 (6.5)</td>
<td>1.00 (0.79, 1.26)</td>
</tr>
<tr>
<td>Other (widowed, separated, divorced)</td>
<td>226 (17.8)</td>
<td>2704 (21.2)</td>
<td>0.80 (0.69, 0.93)</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Income support status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No income support</td>
<td>922 (72.4)</td>
<td>9037 (71.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>Some income support</td>
<td>308 (24.2)</td>
<td>3189 (25.1)</td>
<td>0.95 (0.83, 1.08)</td>
</tr>
<tr>
<td>High income support</td>
<td>42 (3.3)</td>
<td>501 (3.9)</td>
<td>0.82 (0.60, 1.13)</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Residence status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More urban</td>
<td>538 (42.3)</td>
<td>5143 (40.4)</td>
<td>1.00</td>
</tr>
<tr>
<td>More rural</td>
<td>734 (57.7)</td>
<td>7584 (59.6)</td>
<td>0.93 (0.82, 1.04)</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>OC use (&gt;2 years prior to index date)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light users (&lt;12 Rx)</td>
<td>992 (77.9)</td>
<td>10203 (80.2)</td>
<td>1.00</td>
</tr>
<tr>
<td>Heavy users (≥12 Rx)</td>
<td>281 (22.1)</td>
<td>2527 (19.9)</td>
<td>1.15 (1.00, 1.32)</td>
</tr>
<tr>
<td>HT use (&gt;2 years prior to index date)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light users (&lt;12 Rx)</td>
<td>912 (71.6)</td>
<td>9380 (73.7)</td>
<td>1.00</td>
</tr>
<tr>
<td>Heavy users (≥12 Rx)</td>
<td>361 (28.4)</td>
<td>3350 (26.3)</td>
<td>1.12 (0.98, 1.26)</td>
</tr>
<tr>
<td>Age as of index date (yrs)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>60.5 (11.7)</td>
<td>60.4 (11.7)</td>
<td>p = 0.87**</td>
</tr>
<tr>
<td>Age categories</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40 (%)</td>
<td>34 (2.7)</td>
<td>340 (2.7)</td>
<td>1.00</td>
</tr>
<tr>
<td>40-55 (%)</td>
<td>410 (32.2)</td>
<td>4181 (32.8)</td>
<td>1.01 (0.75, 1.35)</td>
</tr>
<tr>
<td>&gt;55 (%)</td>
<td>829 (65.1)</td>
<td>8209 (64.5)</td>
<td>0.77 (0.58, 1.04)</td>
</tr>
<tr>
<td>Length of enrolment (yrs)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>27.9 (3.3)</td>
<td>27.9 (3.4)</td>
<td>p = 0.75**</td>
</tr>
<tr>
<td>Median (range)</td>
<td>29 (10, 30)</td>
<td>29 (10, 30)</td>
<td></td>
</tr>
</tbody>
</table>

**P-values: t-test

Note: Total percentages may exceed 100% due to rounding
SSRI distribution in cases and controls

Table 4 shows the total SSRI exposure distribution among cases and controls during the three exposure windows of interest in this study. Almost 15% of cases and controls had filled at least one SSRI prescription during the study period more than two years prior to index date (14.5% of cases and 14.6% of controls). However, 23% of these cases and 21% of these controls who were SSRI users had only one prescription dispensed during the study period.

For the highest SSRI exposure category of interest during the period of time more than two years prior to index date, 2.4% of cases and 2.3% of controls had filled 36 or more prescriptions. For the period of time 2-7 years prior to index date, 2.0% of cases and 1.6% of controls had filled at least 36 prescriptions. For the exposure window more than 7 years prior to index date, 1.0% of cases and 1.2% of controls had filled at least 24 prescriptions.

Table 4: SSRI distribution for total SSRI use among cases and controls during exposure windows

<table>
<thead>
<tr>
<th>Exposure window</th>
<th>Number of SSRI prescriptions</th>
<th>All participants (N = 14003)</th>
<th>Cases (N = 1273)</th>
<th>Controls (N = 12730)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>No use</td>
<td>11961 (85.4)</td>
<td>1088 (85.5)</td>
<td>10873 (85.4)</td>
<td></td>
</tr>
<tr>
<td>Total use more than 2 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 years prior to index date</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-35 prescriptions</td>
<td>1725 (12.3)</td>
<td>154 (12.1)</td>
<td>1571 (12.3)</td>
<td></td>
</tr>
<tr>
<td>≥ 36</td>
<td>317 (2.3)</td>
<td>31 (2.4)</td>
<td>286 (2.3)</td>
<td></td>
</tr>
<tr>
<td>2-7 years (prior to index date)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No use</td>
<td>12486 (89.2)</td>
<td>1132 (88.9)</td>
<td>11354 (89.2)</td>
<td></td>
</tr>
<tr>
<td>1-35 prescriptions</td>
<td>1291 (9.2)</td>
<td>116 (9.1)</td>
<td>1175 (9.2)</td>
<td></td>
</tr>
<tr>
<td>≥ 36</td>
<td>226 (1.6)</td>
<td>25 (2.0)</td>
<td>201 (1.6)</td>
<td></td>
</tr>
<tr>
<td>&gt;7 years (prior to index date)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No use</td>
<td>12885 (92.0)</td>
<td>1178 (92.5)</td>
<td>11707 (92.0)</td>
<td></td>
</tr>
<tr>
<td>1-23 prescriptions</td>
<td>958 (6.8)</td>
<td>82 (6.4)</td>
<td>876 (6.9)</td>
<td></td>
</tr>
<tr>
<td>≥ 24</td>
<td>160 (1.1)</td>
<td>13 (1.0)</td>
<td>147 (1.2)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Total percentages may exceed 100% due to rounding
4.1.3 Characteristics related to the use of SSRIs among controls

Table 5 summarizes the frequency distribution of selected characteristics related to SSRI use among controls. For this table, controls are divided into three groups based on the number of prescriptions filled over the entire study period (excluding the two year period prior to diagnosis): non SSRI users (‘control’ group for risk analysis), light users consisting of women with 1-35 prescriptions recorded, and thirdly, heavy users consisting of women with 36 or more prescriptions filled. This analysis was undertaken to determine which characteristics might be related to the use of SSRIs among controls (as representative of the underlying population) and therefore potential confounders.

The mean length of enrolment for controls across all three categories of SSRI use (nonusers, light users and heavy users) was similar (p = 0.35). Covariates related to SSRI use included marital status (p< 0.0001), income support status (p< 0.0001), hormone therapy use (p< 0.0001) and age category (p<0.0001). In other words, compared to nonusers, SSRI users (including both light and heavy users) were more likely to be single or widowed/separated/ divorced and to have received income support. Also, as compared to nonusers of SSRIs, a higher percentage of SSRI users were classified as heavy users of HT and SSRI users were more likely to be in the age category 40-55 and less likely to be >55 years of age.

A significant difference in the mean age was observed across the three SSRI exposure categories (p< 0.0001) with users more likely to be younger (mean age for heavy SSRI users was 58.6 years and 60.7 years for nonusers). In summary, among controls in this study population, mean age, marital status, income support status, hormone therapy use and age category were related to SSRI use.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Nonusers* (N = 10873)</th>
<th>Light users (N = 1571)</th>
<th>SSRI users (N = 1857)</th>
<th>p-value**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>1 – 35 Rx n (%)</td>
<td>≥ 36 Rx n (%)</td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>7940 (73.0)</td>
<td>1084 (69.0)</td>
<td>172 (60.1)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>Single</td>
<td>687 (6.3)</td>
<td>108 (6.9)</td>
<td>32 (11.2)</td>
<td></td>
</tr>
<tr>
<td>Other (widowed, separated, divorced)</td>
<td>2243 (20.6)</td>
<td>379 (24.1)</td>
<td>82 (28.7)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income support status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No income support</td>
<td>7850 (72.2)</td>
<td>1019 (64.9)</td>
<td>168 (58.7)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>Some support</td>
<td>2641 (24.3)</td>
<td>460 (29.3)</td>
<td>88 (30.8)</td>
<td></td>
</tr>
<tr>
<td>High support</td>
<td>379 (3.5)</td>
<td>92 (5.9)</td>
<td>30 (10.5)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residence status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More urban</td>
<td>4368 (40.2)</td>
<td>645 (41.1)</td>
<td>130 (45.5)</td>
<td>p = 0.17</td>
</tr>
<tr>
<td>More rural</td>
<td>6502 (59.8)</td>
<td>926 (58.9)</td>
<td>156 (54.5)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light users (&lt;12 Rx)</td>
<td>8745 (80.4)</td>
<td>1234 (78.6)</td>
<td>224 (78.3)</td>
<td>p = 0.16</td>
</tr>
<tr>
<td>Heavy users (≥ 12 Rx)</td>
<td>2128 (19.6)</td>
<td>337 (21.5)</td>
<td>62 (21.7)</td>
<td></td>
</tr>
<tr>
<td>HT use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light users (&lt;12 Rx)</td>
<td>8218 (75.6)</td>
<td>997 (63.5)</td>
<td>165 (57.7)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>Heavy users (≥ 12 Rx)</td>
<td>2655 (24.4)</td>
<td>574 (36.5)</td>
<td>121 (42.3)</td>
<td></td>
</tr>
<tr>
<td>Age as of index date</td>
<td>Mean (SD)</td>
<td>60.7 (11.6)</td>
<td>59.1 (11.9)</td>
<td>58.6 (10.9)</td>
</tr>
<tr>
<td>Age categories</td>
<td>&lt;40 (%)</td>
<td>283 (2.6)</td>
<td>50 (3.2)</td>
<td>7 (2.5)</td>
</tr>
<tr>
<td></td>
<td>40-55 (%)</td>
<td>3481 (32.0)</td>
<td>584 (37.2)</td>
<td>116 (40.6)</td>
</tr>
<tr>
<td></td>
<td>&gt;55 (%)</td>
<td>7109 (65.4)</td>
<td>937 (59.6)</td>
<td>163 (57.0)</td>
</tr>
<tr>
<td>Length of enrolment</td>
<td>Mean (SD)</td>
<td>27.9 (3.3)</td>
<td>27.9 (3.5)</td>
<td>28.1 (3.3)</td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>29 (10, 30)</td>
<td>29 (10, 30)</td>
<td>29 (10,30)</td>
</tr>
</tbody>
</table>

* ‘Control’ group for SSRI users (heavy or light users) is nonuser category

** p-values:
- Chi-square test for categorical variables
- Anova F-test for continuous variables (age and length of enrolment)

**Note:** Total percentages may exceed 100% due to rounding
   Subjects with missing covariate information excluded
4.2 Multivariate analyses

4.2.1 Building a parsimonious model of predictive covariates

As described in detail in section 3.9.3, a backward stepwise elimination procedure was used to select potential confounders for this thesis project. This method of confounder selection was chosen since there are several SSRI exposure representations of interest and it is preferable to adjust for the same covariates in each model. The following covariates remained in a parsimonious model using a p-value of 0.15 as the criterion for retention in the model: marital status, oral contraceptive and hormone therapy use. Age was included in all models. Therefore all models used to examine the relationship between SSRI exposures and breast cancer risk included age, marital status, and OC and HT use as potential confounders.

4.2.2 Multivariate risk analysis

Breast cancer risk by duration of combined SSRI use during each exposure window

Table 6 shows the unadjusted and adjusted odds ratios and 95% confidence intervals for breast cancer according to duration of exposure for all SSRIs combined during each of the three exposure windows of interest: the overall study period (excluding two years prior to index date), 2-7 years and more than seven years prior to index date. Number of prescriptions dispensed during each exposure window was used to represent duration of SSRI use. For the analysis related to overall use during the entire study period two years prior to index date and 2-7 years prior to index date, duration of SSRI use was categorized into three levels: never, 1-35 prescriptions and 36 or more prescriptions dispensed. Due to small sample sizes for highly exposed subjects, exposure categories representing SSRI use for the period of time more than seven years prior to
index date were: never, 1-23 and 24 or more prescriptions dispensed. The referent group for all OR estimates was comprised of subjects with no SSRI prescriptions during the specific exposure window being analyzed.

For the main overall exposure window more than two years prior to index date, results indicate that relative to nonusers, use of SSRIs for a total of three years of more as estimated by 36 or more prescriptions filled was not associated with an increase in breast cancer risk (OR = 1.08, CI = 0.74-1.58). For exposures 2-7 years prior to index date, no association was observed for subjects with 36 or more prescriptions filled compared to nonusers during this time period (OR = 1.25, CI = 0.82-1.90). Similarly, for the period of time more than seven years prior to index date, no suggestion of increased risk was found for the highly exposed subjects (24 or more prescriptions filled) as compared to nonusers during this time period (OR = 0.87, CI = 0.49-1.55).

Table 6: Odds ratios (OR) (crude and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of SSRI prescriptions during each exposure window

<table>
<thead>
<tr>
<th>Exposure window</th>
<th>Number of SSRI prescriptions</th>
<th>Cases N=1273</th>
<th>Controls N=12730</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total use more than 2 years prior to index date</td>
<td>No use*</td>
<td>1088</td>
<td>10873</td>
<td>1.00 (0.98, 1.17)</td>
<td>1.00 (0.98, 1.17)</td>
</tr>
<tr>
<td></td>
<td>1-35</td>
<td>154</td>
<td>1571</td>
<td>0.98 (0.82, 1.17)</td>
<td>1.08 (0.74, 1.58)</td>
</tr>
<tr>
<td></td>
<td>≥ 36</td>
<td>31</td>
<td>286</td>
<td>1.08 (0.74, 1.58)</td>
<td>1.08 (0.74, 1.58)</td>
</tr>
<tr>
<td>2-7 years (prior to index date)</td>
<td>No use*</td>
<td>1132</td>
<td>11354</td>
<td>1.00 (0.99, 1.21)</td>
<td>1.00 (0.99, 1.21)</td>
</tr>
<tr>
<td></td>
<td>1-35</td>
<td>116</td>
<td>1175</td>
<td>0.99 (0.81, 1.21)</td>
<td>1.25 (0.82, 1.90)</td>
</tr>
<tr>
<td></td>
<td>≥ 36</td>
<td>25</td>
<td>201</td>
<td>1.25 (0.82, 1.90)</td>
<td>1.25 (0.82, 1.90)</td>
</tr>
<tr>
<td>&gt; 7 years (prior to index date)</td>
<td>No use*</td>
<td>1178</td>
<td>11707</td>
<td>1.00 (0.93, 1.18)</td>
<td>0.93 (0.74, 1.18)</td>
</tr>
<tr>
<td></td>
<td>1-23</td>
<td>82</td>
<td>876</td>
<td>0.93 (0.74, 1.18)</td>
<td>0.93 (0.74, 1.18)</td>
</tr>
<tr>
<td></td>
<td>≥ 24</td>
<td>13</td>
<td>147</td>
<td>0.88 (0.50, 1.56)</td>
<td>0.87 (0.49, 1.55)</td>
</tr>
</tbody>
</table>

* Referent group: ‘No use’ refers to subjects with no SSRI prescriptions during the specified time period: i.e. subject may have had SSRIs during other time periods
** Odds ratios adjusted for age, marital status, OC and HT use
Breast cancer risk by total dosage of combined SSRI use during each exposure window

Table 7 shows the unadjusted and adjusted odds ratios and 95% confidence intervals for breast cancer according to total dosage of SSRI exposure for all SSRIs combined during each of the three exposure windows of interest. As described in Chapter 3, total dosage for each SSRI was calculated using fluoxetine-equivalent dosages. For each exposure window, dosage of SSRI use was categorized into three levels of exposure based on the control distribution of SSRI use with 10% of exposed controls being in the highest exposure category. The referent category for all OR estimates was comprised of subjects with no SSRI prescriptions during the specific exposure window being analyzed.

For the main overall exposure window more than two years prior to index date, no association was observed between breast cancer and dosage of SSRI use (highest dosage category: \( \geq 38,680 \) mg) relative to nonusers (OR = 1.02, CI = 0.63-1.64). Similarly, relative to nonusers, there was no evidence of increased risk associated with the highest dosage category of SSRI use 2-7 years (OR=0.73, CI=0.38-1.39) and more than seven years (OR=1.06, CI=0.57-2.00) prior to index date.
Table 7: Odds ratios (OR) (crude and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total SSRI dosage*** during each exposure window

<table>
<thead>
<tr>
<th>Exposure time period</th>
<th>Total SSRI dosage (mg) ***</th>
<th>Cases</th>
<th>Controls</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=1273</td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total use more than 2 years prior to index date</td>
<td>No use*</td>
<td>1088 (85.5)</td>
<td>10873 (85.4)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1-38,679 mg</td>
<td>166 (13.0)</td>
<td>1671 (13.1)</td>
<td>0.99 (0.84, 1.18)</td>
<td>0.99 (0.83, 1.18)</td>
</tr>
<tr>
<td></td>
<td>≥ 38,680 mg</td>
<td>19 (1.5)</td>
<td>186 (1.5)</td>
<td>1.02 (0.63, 1.64)</td>
<td>1.02 (0.63, 1.64)</td>
</tr>
<tr>
<td>2-7 years (prior to index date)</td>
<td>No use*</td>
<td>1132 (88.9)</td>
<td>11354 (89.2)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1-32,319 mg</td>
<td>131 (10.3)</td>
<td>1238 (9.7)</td>
<td>1.06 (0.88, 1.28)</td>
<td>1.06 (0.87, 1.28)</td>
</tr>
<tr>
<td></td>
<td>≥ 32,320 mg</td>
<td>10 (0.8)</td>
<td>138 (1.1)</td>
<td>0.73 (0.38, 1.39)</td>
<td>0.73 (0.38, 1.39)</td>
</tr>
<tr>
<td>&gt;7 years (prior to index date)</td>
<td>No use*</td>
<td>1178 (92.5)</td>
<td>11707 (92.0)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1-21,629 mg</td>
<td>84 (6.6)</td>
<td>920 (7.2)</td>
<td>0.91 (0.72,1.14)</td>
<td>0.91 (0.72,1.14)</td>
</tr>
<tr>
<td></td>
<td>≥ 21,630 mg</td>
<td>11 (0.9)</td>
<td>103 (0.8)</td>
<td>1.06 (0.57,1.98)</td>
<td>1.06 (0.57, 2.00)</td>
</tr>
</tbody>
</table>

*Referent group: ‘No use’ refers to subjects with no SSRI prescriptions during the specified time period: i.e. subject may have had SSRI prescriptions during other time periods

** Odds ratios adjusted for age, marital status, OC and HT use

*** Combined dosage calculated using fluoxetine-equivalent dosages (Defined daily dose (DDD) methodology, WHO) for each individual SSRI prescription whereby 20 mg of fluoxetine = 20 mg citalopram = 20 mg paroxetine = 50 mg of sertraline = 100 mg of fluvoxamine

Breast cancer risk by duration of individual SSRI use during main exposure window

Evaluation of risk associated with duration of exposure to each individual SSRI during the main overall exposure window more than two years prior to index date is presented in Tables 8-12. Duration of use was represented as total number of prescriptions filled. As described previously, for each SSRI drug, ‘any’ use referred to women who had used the SSRI of interest alone or in combination with one or more other SSRIs. ‘Exclusive’ users represented a sub-set of ‘any’ users who had only used one type of SSRI during the study period. For all OR estimates, nonusers of SSRIs during the
main exposure window more than two years prior to index date served as the referent group.

Overall, as these ORs indicate, there is no evidence of increased risk of breast cancer for highly exposed users of each individual SSRI (including ‘any’ or ‘exclusive’ use) relative to nonusers of SSRIs. The highest ORs were observed for ‘exclusive’ use of paroxetine (OR = 1.42, CI = 0.67-2.98) and sertraline (OR = 2.22, CI = 0.84-5.86), and ‘any’ use of citalopram (OR = 3.23, CI = 0.87-11.99) but sample sizes were small and the confidence intervals were wide.

Small sample sizes did not allow a complete evaluation of risk associated with the use of citalopram. In addition, small sample sizes of SSRI users of individual SSRIs did not allow an evaluation of risk associated with use during the other two exposure windows of interest (2-7 years and more than seven years prior to index date).

**Table 8: Paroxetine: Odds ratios (OR) (unadjusted and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of prescriptions**

<table>
<thead>
<tr>
<th>Total number of prescriptions</th>
<th>Cases N=1273</th>
<th>Controls N=12730</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No SSRI use*</td>
<td>1088</td>
<td>10873</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>ANY use of paroxetine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>64</td>
<td>632</td>
<td>1.01 (0.78, 1.32)</td>
<td>1.00 (0.77, 1.32)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>14</td>
<td>119</td>
<td>1.18 (0.67, 2.05)</td>
<td>1.14 (0.65, 2.00)</td>
</tr>
<tr>
<td>Exclusive users***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>39</td>
<td>341</td>
<td>1.14 (0.82, 1.60)</td>
<td>1.13 (0.80, 1.58)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>8</td>
<td>55</td>
<td>1.46 (0.69, 3.06)</td>
<td>1.42 (0.67, 2.98)</td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during the main exposure window
** Odds ratios adjusted for age, marital status, OC and HT use
***Exclusive users: subset of paroxetine users that only took paroxetine
Table 9: Sertraline: Odds ratios (OR) (unadjusted and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of prescriptions

<table>
<thead>
<tr>
<th>Total number of prescriptions</th>
<th>Cases N=1273</th>
<th>Controls N=12730</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No SSRI use*</td>
<td>1088</td>
<td>10873</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>ANY use of sertraline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>31</td>
<td>389</td>
<td>0.80 (0.55, 1.15)</td>
<td>0.79 (0.55, 1.15)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>6</td>
<td>66</td>
<td>0.91 (0.39, 2.10)</td>
<td>0.92 (0.40, 2.14)</td>
</tr>
<tr>
<td>Exclusive users***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>18</td>
<td>180</td>
<td>1.00 (0.61, 1.63)</td>
<td>1.00 (0.61, 1.63)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>5</td>
<td>23</td>
<td>2.17 (0.82, 5.73)</td>
<td>2.22 (0.84, 5.86)</td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during main exposure window
** Odds ratios adjusted for age, marital status, OC and HT use
*** Exclusive users: subset of sertraline users that only took sertraline

Table 10: Fluoxetine: Odds ratios (OR) (unadjusted and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of prescriptions

<table>
<thead>
<tr>
<th>Total number of prescriptions</th>
<th>Cases N=1273</th>
<th>Controls N=12730</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No SSRI use*</td>
<td>1088</td>
<td>10873</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>ANY use of fluoxetine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>62</td>
<td>662</td>
<td>0.94 (0.72, 1.22)</td>
<td>0.94 (0.71, 1.22)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>17</td>
<td>164</td>
<td>1.04 (0.63, 1.71)</td>
<td>1.03 (0.62, 1.71)</td>
</tr>
<tr>
<td>Exclusive users***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>30</td>
<td>359</td>
<td>0.84 (0.57, 1.22)</td>
<td>0.83 (0.57, 1.22)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>9</td>
<td>94</td>
<td>0.96 (0.48, 1.90)</td>
<td>0.96 (0.48, 1.91)</td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during main exposure window
** Odds ratios adjusted for age, marital status, OC and HT use
*** Exclusive users: subset of fluoxetine users that only took fluoxetine
### Table 11: Fluvoxamine: Odds ratios (OR) (unadjusted and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of prescriptions

<table>
<thead>
<tr>
<th>Total number of prescriptions</th>
<th>Cases N=1273</th>
<th>Controls N=12730</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No SSRI use*</td>
<td>1088</td>
<td>10873</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>ANY use of fluvoxamine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>30</td>
<td>312</td>
<td>0.96 (0.66, 1.41)</td>
<td>0.97 (0.66, 1.41)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>3</td>
<td>44</td>
<td>0.68 (0.21, 2.20)</td>
<td>0.69 (0.21, 2.24)</td>
</tr>
<tr>
<td>Exclusive users***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>15</td>
<td>131</td>
<td>1.14 (0.67, 1.96)</td>
<td>1.15 (0.67, 1.96)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>3</td>
<td>26</td>
<td>1.15 (0.35, 3.82)</td>
<td>1.15 (0.35, 3.82)</td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during main exposure window  
** Odds ratios adjusted for age, marital status, OC and HT use  
*** Exclusive users: subset of fluvoxamine users that only took fluvoxamine

### Table 12: Citalopram: Odds ratios (OR) (unadjusted and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of prescriptions

<table>
<thead>
<tr>
<th>Total number of prescriptions</th>
<th>Cases N=1273</th>
<th>Controls N=12730</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No SSRI use*</td>
<td>1088</td>
<td>10873</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>ANY use of citalopram</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>13</td>
<td>189</td>
<td>0.69 (0.39, 1.21)</td>
<td>0.69 (0.39, 1.21)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>3</td>
<td>9</td>
<td>3.34 (0.90, 12.33)</td>
<td>3.23 (0.87, 11.99)</td>
</tr>
<tr>
<td>Exclusive users***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>5</td>
<td>80</td>
<td>0.63 (0.25, 1.55)</td>
<td>0.63 (0.26, 1.56)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>0</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during main exposure window  
** Odds ratios adjusted for age, marital status, OC and HT use  
*** Exclusive users: subset of citalopram users that only took citalopram
Breast cancer risk by duration of use of ‘high inhibitors’ of serotonin uptake

Table 13 shows results of a sub-analysis done to test the hypothesis that the use of any or all of the high inhibitors of serotonin reuptake (namely paroxetine, sertraline and fluoxetine) may be associated with an increased risk of breast cancer. The referent category was comprised of women who had never used SSRIs during the main exposure window of interest, that is, during the period of time more than two years prior to index date. Relative to nonusers of SSRIs, no suggestion of increased risk was observed for the highly exposed group of paroxetine, sertraline and/or fluoxetine users (adjusted OR = 1.04, CI = 0.74-1.47).

Table 13: Odds ratios (OR) (unadjusted and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of prescriptions for paroxetine, sertraline and/or fluoxetine

<table>
<thead>
<tr>
<th>Total number of prescriptions</th>
<th>Cases</th>
<th>Controls</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No SSRI use*</td>
<td>1088</td>
<td>10873</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1-23</td>
<td>124</td>
<td>1256</td>
<td>0.99 (0.81, 1.20)</td>
<td>0.99 (0.81, 1.20)</td>
</tr>
<tr>
<td>≥ 24</td>
<td>37</td>
<td>354</td>
<td>1.05 (0.74, 1.48)</td>
<td>1.04 (0.74, 1.47)</td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during main exposure window
** Odds ratios adjusted for age, marital status, OC and HT use

Breast cancer risk by duration of SSRI exposure and age at diagnosis

For this analysis, age at index date was used as a proxy for menopausal status with pre-menopausal defined as 55 years of age or younger and post-menopausal defined as over age 55. The significance of an interaction term (age at index date*SSRI exposure) was tested in order to evaluate whether menopausal status at index date modified the association between SSRI use and breast cancer risk. For this analysis, the exposure
contrast of interest was nonusers versus ‘heavy’ users (36 or more prescriptions filled). Results indicated that the interaction term was not significant (P-value interaction = 0.76).

Given the evidence suggesting that SSRIs play a role in breast cancer development through hormonal (prolactin) pathways and that the analysis may have lacked power to detect the interaction, further stratified analysis (see Tables 14 and 15) was done to assess differences in risk between the two subgroups of women defined by age at index date. No increase in risk was observed among the highly exposed (36 or more prescriptions during the main exposure window more than two years prior to index date) relative to nonusers of SSRIs in the 55 years of age or younger age group (OR = 0.91, CI = 0.49-1.70). For women over age 55, relative to nonusers, heavy SSRI use was also not associated with an increase in breast cancer risk (OR = 1.21, CI = 0.76-1.94).

Table 14: Odds ratios (OR) (crude and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of SSRI prescriptions during main exposure window for subjects 55 years of age and younger

<table>
<thead>
<tr>
<th>Number of SSRI prescriptions</th>
<th>Cases N=444</th>
<th>Controls N= 4521</th>
<th>Unadjusted OR 95% CI</th>
<th>Adjusted OR** 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No use*</td>
<td>371 (83.6)</td>
<td>3764 (83.3)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1-35</td>
<td>62 (14.0)</td>
<td>634 (14.0)</td>
<td>1.00 (0.75, 1.32)</td>
<td>0.99 (0.75, 1.32)</td>
</tr>
<tr>
<td>≥36</td>
<td>11 (2.5)</td>
<td>123 (2.7)</td>
<td>0.93 (0.50, 1.75)</td>
<td>0.91 (0.49, 1.70)</td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during the main exposure window  
** Odds ratios adjusted for age, marital status, OC and HT use
Table 15: Odds ratios (OR) (crude and multivariate-adjusted) and 95% confidence intervals (CI) for breast cancer according to total number of SSRI prescriptions during main exposure window for subjects over age 55

<table>
<thead>
<tr>
<th>Number of SSRI prescriptions</th>
<th>Cases N=829</th>
<th>Controls N=8209</th>
<th>Unadjusted OR</th>
<th>Adjusted OR**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No use*</td>
<td>717 (86.5)</td>
<td>7109 (86.6)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1-35</td>
<td>92 (11.1)</td>
<td>937 (11.4)</td>
<td>0.97 (0.78, 1.22)</td>
<td>0.98 (0.78, 1.23)</td>
</tr>
<tr>
<td>≥36</td>
<td>20 (2.4)</td>
<td>163 (2.0)</td>
<td>1.22 (0.76, 1.95)</td>
<td>1.21 (0.76, 1.94)</td>
</tr>
</tbody>
</table>

* Referent group: No SSRI use during the main exposure window
** Odds ratios adjusted for age, marital status, OC and HT use

**Rothman confounder assessment**

As previously described, an alternative analysis (shown in Appendix 4, Table 4.1) was done to identify potential confounders using the change-in-estimate method (Greenland and Rothman, 1998, p.257). Confounders were identified using this method based on a 10% change in the odds ratio associated with the exposure contrast of interest (nonusers versus those with ≥ 36 prescriptions dispensed).

Using this approach, all of the potential confounders were removed from the model. The OR for the fully adjusted model (including all potential confounders) was 1.09 (CI = 0.75-1.59) and the OR (excluding all potential confounders) was 1.08 (CI = 0.74-1.58). The percent change in risk estimate (OR) between these two models was 0.5%.

Comparing nonusers with those having 36 or more prescriptions, the OR and 95% confidence intervals for the model with all potential confounders removed built using the change-in-risk estimate strategy was identical to the adjusted OR observed in the model that controlled for age, marital status, OC and HT use.
4.3 Stage two subject sampling method

For the larger two-stage study, a sub-set of stage one cases and controls is being selected for the purposes of collecting confounder information by telephone interview. To achieve objective three of this thesis project, the first sample of cases and controls was selected according to both disease and exposure status with the intent to use a ‘balanced design’ approach whereby approximately equal numbers of cases and controls would be selected from each of three main SSRI exposure categories (nonusers, 1-35 prescriptions and \( \geq 36 \) prescriptions filled during the period of time more than two years prior to index date). In order to avoid potential long delays between breast cancer diagnosis and interviews, this selection of stage two cases and controls was limited to subjects with an index date in 2004 and 2005. Table 16 summarizes the exposure distribution for the 858 cases and 8592 controls with index dates in 2004 or 2005.

Table 16: Case-control exposure distribution (index date 2004-2005)

<table>
<thead>
<tr>
<th>Number of SSRI prescriptions</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>No use</td>
<td>727 (85.0%)</td>
<td>7298 (85.0%)</td>
</tr>
<tr>
<td>1-35</td>
<td>105 (12.0%)</td>
<td>1074 (12.5%)</td>
</tr>
<tr>
<td>( \geq 36 )</td>
<td>26 (3.0%)</td>
<td>220 (2.5%)</td>
</tr>
<tr>
<td>Totals</td>
<td>858 (100%)</td>
<td>8592 (100%)</td>
</tr>
</tbody>
</table>

4.3.1 Stage two sample invited to participate in interviews

In order to select the stage two sample to be invited for interviews from the subjects in Table 16, several steps were necessary. First, it was necessary to estimate the total number of subjects with completed interviews required for stage two of the larger study. Since sample size calculations indicated that the final stage two sample (based on
a total of four years of subject accrual from 2004-2007) required 200 cases and 200 controls with completed interviews, it was determined that this first selection process (based on two years of data: 2004-2005 data only) would require a minimum of 100 cases and 100 controls with completed interviews. In addition, response rates from previous studies were considered in order to estimate the number of subjects that needed to be invited for interviews in order to have the required number of subjects with completed interviews in the final stage two sample. Given an anticipated 60% response rate for cases and a 40% response rate for controls, the estimated sample of subjects with index dates in 2004 and 2005 invited for the interviews would therefore have to be inflated from 100 cases and 100 controls to an estimated 174 cases and 261 controls.

Ideally, for the ‘balanced design’ sampling approach, equal numbers of cases and controls would be invited to the interviews from each of the three exposure categories by over-sampling of the cells with small numbers of subjects (especially highly exposed cases) and sampling cells with larger numbers less frequently (unexposed controls). However, in this study population, few exposed cases (especially highly exposed cases) meant that it was not possible to have equal numbers of cases in each of the three exposure categories (no use, 1-35 and \( \geq 36 \) prescriptions dispensed). Therefore, in order to increase the likelihood that the final analysis using the stage two data would include enough exposed cases, 100% of cases from both of the SSRI exposure categories (1-35 and \( \geq 36 \) prescriptions dispensed) were invited to participate in the interviews. Table 17 shows the estimated required number of stage two subjects with index date in 2004 or 2005 from each exposure category that needed to be invited for the interviews taking into account anticipated response rates. In Table 17, for controls, the ‘balanced’ design
approach was used, whereby an equal number of controls would be selected from each exposure category.

Table 17: *Estimated* required number of stage two cases and controls that need to be *invited* for interviews considering expected response rates

<table>
<thead>
<tr>
<th>Number of SSRI prescriptions</th>
<th>Cases (% of subjects from table 16)</th>
<th>Controls (% of subjects from table 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No use</td>
<td>58 (8.0%)</td>
<td>87 (1.2%)</td>
</tr>
<tr>
<td>1-35</td>
<td>105 (100%)</td>
<td>87 (8.1%)</td>
</tr>
<tr>
<td>≥ 36</td>
<td>26 (100%)</td>
<td>87 (39.5%)</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>174</td>
<td>261</td>
</tr>
</tbody>
</table>

For the next step, the Ranuni function in SAS was used to generate a random number between 0 and 1 for each study subject with an index date in 2004 or 2005. As shown in Table 18, cases and controls were randomly selected using this random number whereby 100% of cases in both of the SSRI user exposure categories (1-35 and ≥36 prescriptions dispensed) and 8.3% of unexposed cases (nonusers) were invited to participate in the interview process. Rather than inviting equal numbers of controls from each of the exposed categories as shown in Table 17, it was decided to sample more frequently from the SSRI users in order to increase the likelihood of having an adequate number of exposed controls in the stage two sample. Therefore, 45% of controls were randomly selected from the highest exposure category, 8.6% from the 1-35 prescriptions dispensed group and 1.1% of unexposed controls were randomly selected to be invited for the stage two interviews.

Table 18 shows the exposure distribution for the actual final *invited sample* of cases and controls randomly selected from the 2004-2005 case-control subjects using the Ranuni procedure and utilizing the percentages described in the above text. Study
identification numbers of these 191 cases and 268 controls were provided to the co-investigator at SK Health. Invitations for interviews have been sent out to each of these subjects and interviews are presently taking place.

Table 18: Actual final sample of cases and controls invited for interviews

<table>
<thead>
<tr>
<th>Number of SSRI prescriptions</th>
<th>Cases (% of subjects from table 16)</th>
<th>Controls (% of subjects from table 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No use</td>
<td>60 (8.3%)</td>
<td>77 (1.1%)</td>
</tr>
<tr>
<td>1-35</td>
<td>105 (100%)</td>
<td>92 (8.6%)</td>
</tr>
<tr>
<td>≥ 36</td>
<td>26 (100%)</td>
<td>99 (45.0%)</td>
</tr>
<tr>
<td>Totals</td>
<td>191</td>
<td>268</td>
</tr>
</tbody>
</table>

4.3.2 Anticipated stage two sample with completed interviews

Table 19 shows the anticipated case-control sample with completed interviews assuming that response rates are as predicted across each exposure category (60% for cases and 40% for controls). This sample of cases and controls chosen from the stage one sample with index dates from 2004-2005 represents approximately one-half of the final stage two subject sample. The other 50% of cases and controls will be selected from subjects with index dates from 2006-2007 for an expected final total of about 200 cases and 200 controls with completed interviews.

Table 19: Anticipated stage two sample (with 2004-2005 index dates) with completed interviews assuming response rates are similar for each SSRI exposure category (60% for cases and 40% for controls)

<table>
<thead>
<tr>
<th>Number of SSRI prescriptions</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>No use</td>
<td>36</td>
<td>31</td>
</tr>
<tr>
<td>1-35</td>
<td>63</td>
<td>37</td>
</tr>
<tr>
<td>≥ 36</td>
<td>16</td>
<td>40</td>
</tr>
<tr>
<td>Totals</td>
<td>115</td>
<td>108</td>
</tr>
</tbody>
</table>
Chapter 5.0 Discussion

5.1 Summary of main findings

The primary objective of this thesis project was to determine breast cancer risk associated with dosage, duration and timing of SSRI use, controlling for a limited set of potential confounders. This large population-based nested case-control linkage study incorporated a prescription database approach to ascertaining exposure to SSRIs, and risk was examined during specific exposure windows within the context of a biologically plausible conceptual model of breast cancer development. Overall, results do not support a relationship between breast cancer and long-term use of SSRIs combined during specific time periods (total use more than two years prior to index date, 2-7 years and >7 years prior to index date). In addition, no suggestion of increased risk of breast cancer was detected in relation to individual SSRI exposures.

In this chapter, results will be summarized and interpreted within the context of relevant previous epidemiologic research, and the strengths and limitations of this thesis research. Suggestions for future research will be followed by a brief discussion of the overall contribution of this thesis project.

5.2 Risk associated with long-term combined SSRI use

It is difficult to compare results of this study directly with those of previous studies because of variation in SSRI exposure measurement methods (self-reports versus prescription database) and differences in measures used to define long-term duration and dosage of SSRI use. Also, few studies assessed risk associated with long-term use of comparable duration to this thesis project (that is, the equivalent of three or more years of
use as defined by 36 or more prescriptions dispensed) or risk associated with SSRI exposures during specific time periods.

In terms of design (population-based, nested case-control study) and exposure characterization (number of prescriptions dispensed used to quantify SSRI use), this thesis project is most similar to the study by Fulton-Kehoe et al. (2006) in which no increase in breast cancer risk associated with long-term use (defined as 21 or more prescriptions dispensed) was observed.

Despite variations in SSRI exposure definitions used to quantify long-term duration and dosage of exposures, the results of this thesis are also consistent with most other observational studies that assessed long-term use, including four studies that used self-reported SSRI exposure information (Kelly et al., 1999; Cotterchio et al., 2000; Chien et al., 2006; Coogan et al., 2005), and two studies that ascertained SSRI use from a prescription database (González-Pérez and Rodríguez, 2005; Wang et al., 2001). However, only two of these studies (González-Pérez and Rodríguez, 2005; Coogan et al., 2005) had sufficient sample size to evaluate risk associated with three or more years of use and only one study (González-Pérez and Rodríguez, 2005) quantified total long-term SSRI exposures using a dosage variable. These thesis results are also consistent with Mooreman et al. (2003) despite the author’s interpretation of a ‘modest’ positive association among women who had used SSRIs for 3 years or longer (OR = 2.2, CI = 0.8-6.3).

The effect of timing of SSRI exposures on breast cancer risk was previously investigated in a PhD thesis project by Tamim (2003) (results not published, personal communication, 2007). Using the SK prescription database for exposure information, risk associated with SSRI exposures during two periods of time (2-7 and 6-10 years prior
to index date) was assessed. For the majority of exposure indices, no statistically
significant increase in breast cancer risk was detected. SSRI exposure variables included
ever/never use, and cumulative dosage and average daily dosage both expressed in moles.
Estimation of duration of use was calculated by dividing each time period into three
month intervals and then expressing duration as a percentage of the number of intervals
with at least one prescription. Despite major differences in defining SSRI exposures and
varying time intervals of interest, the results of Tamim (2003) are consistent with this
thesis project.

5.3 Risk associated with the use of individual SSRIs

For specific SSRIs, thesis results pertaining to risk associated with long-term
duration and dosage of use are consistent with most other observational studies that
specifically assessed these relationships (Fulton-Kehoe et al., 2006; Coogan et al., 2005;
González-Pérez and Rodríguez, 2005; Chien et al., 2006; Haque et al., 2005). However,
only one previous study by Coogan et al. (2006) assessed risk associated with individual
SSRI use of equal or longer duration than this thesis study reporting no increased risk
with the use of fluoxetine or sertraline for 2-4 and ≥ 4 years duration of use. Fulton-
Kehoe et al. (2006), using number of prescriptions as a measure of exposure, reported no
increased risk associated with any use of paroxetine, fluoxetine or sertraline (as defined
by 11 or more prescriptions dispensed). Previous case-control studies have reported
marginally increased risk associated with any use of paroxetine (Cotterchio et al., 2000;
Steingart et al., 2003) and sertraline (Steingart et al., 2003). The highest ORs observed in
this study were for ‘exclusive’ users of paroxetine, sertraline and users of ‘any’
citalopram but sample sizes were small and confidence intervals were wide. However,
direct comparisons between these thesis results and previous studies are difficult because of different SSRI exposure definitions used.

### 5.4 Risk and Menopausal status

To evaluate a potentially different effect of SSRI exposure in pre- and post-menopausal women, the association between breast cancer and SSRI use among two subgroups of women was assessed using age at index date as a proxy for menopausal status (pre-menopausal: \( \leq 55 \) and post-menopausal: \( >55 \) years of age). No increased risk was observed among these 2 subgroups, consistent with the two previous studies that examined these relationships (González-Pérez and Rodríguez, 2005; Coogan et al., 2006).

### 5.5 Methodological Considerations

In order to assess the validity of the study results, it is important to consider the strengths and limitations of this study.

#### 5.5.1 Strengths

This study has several strengths. The Saskatchewan databases are recognized as valuable resources for drug utilization review and pharmacoepidemiologic studies (Downey et al., 2005, p.306; Tilson, 1985; Malcom et al., 1993; Rawson et al., 1992) with high levels of quality control, internal auditing and cross database integrity (Tilson, 1985). The SPDP and the SCA cancer registry are population-based and data are available for an extensive period of time (since early 1970’s). Each database has its own internal verification process to ensure identification and demographic accuracy (Downey et al., 2005, p.303). During the linkage process that involves matching on each subject’s
unique health services number (HSN), inconsistencies are manually checked and questionable matches are rejected, minimizing duplicates and mismatches and ensuring a high degree of confidence in the linkage process.

The completeness and accuracy of the breast cancer data are considered to be very high. The SCA cancer registry has attained Silver Standard certification for the last number of years from the North American Association of Central Cancer Registries (NAACCR), a professional organization that develops and promotes uniform data standards for cancer registration (NAACCR website).

The advantages of using a prescription database compared to self-reported SSRI use for exposure information included relatively easy access to large sample sizes and elimination of exposure recall error and bias (Downey, 2005, p.296; West et al., 1997; West et al., 1995). In addition, with the use of detailed prescription database exposure histories, more precise SSRI exposure measures could be created to represent duration and dosage of long-term exposures, and the temporal effects of chronic long-term SSRI use could be assessed more accurately.

Lack of selection bias is also a major strength of this study. There is virtually complete case ascertainment, and controls for this study have a known and equal probability of being randomly selected from the same source population as the cases during the same time interval.

5.5.2 Limitations

Limitations of this study include sources of information bias, confounding and detection bias. The use of the exposure window approach may also bias results.
**Information bias:** There are three potential sources of information bias: medication non-compliance, incomplete SSRI exposure data, and imprecise definition of drug exposure.

(a) **Non-compliance with dispensed medication:** Non-compliance (i.e. subject did not take some or all of prescribed medication) is a possible source of exposure misclassification since it is not known how much of the dispensed SSRI prescription was actually consumed. Non-compliance is a potential limitation of all studies looking at this association, whether SSRI exposure is ascertained from self-reports or prescription databases. Medication noncompliance is a major problem in the treatment of psychiatric disorders (Schwarz, 2006) with estimates for depressive disorders ranging from 10-60% (Lingam, 2002). In a prospective observational study using electronic drug exposure monitors to evaluate day-to-day compliance during three months of follow-up among 69 ‘chronic’ users of SSRIs recruited by pharmacists (‘chronic’ defined as under treatment for an average of 2.5 years), Meijer et al. (2001) reported that 13% (9/69) of patients had at least one lapse of three or more consecutive days (range three to five days). These results however only address compliance over a three-month period and would not necessarily apply to SSRI users over the course of three or more years. Also, sample size was small and there was no information about patients who refused to participate in the study. In addition, one could speculate from a biologic perspective that lapses of this short duration (three to five days) would not influence a potential carcinogenic effect of SSRIs on breast cancer development.

In this study assessing risk associated mainly with long-term SSRI exposures, it was anticipated that misclassification due to non-compliance (primarily overestimation of exposure) would be more likely to occur within the low exposure categories (one or two prescriptions filled), rather than in the contrasts of most interest (e.g. ≥36 prescriptions
versus no exposure). That is, if prescriptions were repeatedly refilled, it was more likely that they were actually being consumed. Therefore, especially within the higher exposure group, misclassification of nonusers of SSRIs as users would be less likely. Further, non-compliance with dispensed medications would likely affect cases and controls equally in this study, leading to non-differential misclassification tending to bias the results towards the null.

(b) Incomplete SSRI exposure data: The SPDP database lacked information on SSRI medications dispensed in hospitals or as samples from physicians (Downey et al., 2005, p.297). Also, escitalopram, the newest SSRI to become available in Canada (since late 2005), was not included in the SK Drug Formulary at the time of data collection for this study and therefore prescriptions filled for this SSRI were not captured in the SPDP prescription database. However, it was anticipated that the degree of potential misclassification (underestimation of exposure) associated with these additional sources of information bias would be uniform between the case and control groups (non-differential misclassification) tending to bias results towards the null.

(c) Imprecise definition of SSRI exposure: Since the SK database does not include information on the prescribed daily dosage regime (pills per dose and total number of doses per day), total number of prescriptions dispensed was used to estimate duration of exposure based on the assumption that each prescription represented a 34-day supply. As was mentioned previously, most drugs in SK (including SSRIs) are typically dispensed to cover a 34-day period based on a contractual agreement between the SK Drug plan and pharmacies, whereby the pharmacist is entitled to charge a dispensing fee for each 34-day supply. However, the SK agreement does not prohibit the pharmacist from dispensing more than a 34-day supply for one fee. Therefore, each prescription recorded in the
database may not have represented the same amount (dosage) of medication, leading to incorrect estimations of exposure (primarily underestimation of exposure) since total dosage of each prescription and number of days between refills was not considered.

Further analysis was done to assess total dosage amounts for each individual SSRI prescription in the case and control groups separately. For fluoxetine, paroxetine and citalopram, cases and controls tended to have the same range of dosages for each prescription dispensed (cases and controls had approximately equal number of prescriptions for high and low dosage amounts dispensed for each SSRI). However, controls tended to have a higher dosage for fluvoxamine and sertraline prescriptions dispensed as compared to cases. Therefore, for controls, individual (fluvoxamine and sertraline) and combined SSRI exposures may have been underestimated, since number of prescriptions dispensed does not take total dosage into account, which would bias these results from the null (overestimation of effect).

Also, to indirectly test the extent of misclassification related to estimating exposure using number of prescriptions, risk was estimated for total dosage of combined SSRI use, and none was detected. Therefore, this evidence confirms that exposure classification errors in relation to duration of use (as estimated by the total number of prescriptions for all SSRIs combined) were unlikely.

**Bias due to confounding:** Risk estimates may have been affected by three sources of confounding (residual, uncontrolled and confounding by indication).

(a) **Sources of residual confounding:** For this study, two indicators of socioeconomic status (SES) were available from the SK databases (income support status and urban/rural status). The nature of the associations between breast cancer risk, SSRI use and SES is
unclear. Evidence suggests that individuals of higher SES (whites with private drug coverage) are more likely to be prescribed an SSRI as compared to a non-SSRI antidepressant (Melfi et al., 2000; Sleath & Shih, 2003; Olfson et al., 1998). Since information on other potential confounding variables related to SES, such as income, occupation and education level are not available in the SK databases, residual confounding may have affected the risk estimates.

In this study population, SSRI users were more likely to receive income support, suggestive that individuals of lower SES were more likely to have used SSRIs and conversely, that individuals of higher SES (those who did not receive income support) were less likely to have used SSRIs. Therefore, given that higher SES is associated with increased breast cancer risk (Hulka and Mooreman, 2001) it would appear that residual confounding due to SES may be less likely in this study population.

In addition, the analysis may not have adequately controlled for HT and OC use since the broad categories chosen to represent exposure to HT and OCs (<12 versus \( \geq 12 \) prescriptions dispensed more than two years prior to index date) did not account for differences in potency of formulations used and may not have represented the most relevant time window of exposure.

In the final analysis for the complete two-stage study, other SES variables will include income and education level. In addition HT and OC use will be represented by variables that consider potency of the formulations and the timing of use.

(b) Sources of uncontrolled confounding: As this study relies solely on SK databases for subject data, the available information on potential confounders is limited. The SK databases did not contain information about established protective effects and risk factors for breast cancer that may have distorted (confounded) the risk estimates, such as
information related to medical and reproductive history, family history of cancer, anthropometric variables (height and weight), and lifestyle factors including smoking, physical activity and alcohol consumption. If any of these factors were associated with SSRI use and were present to a greater or lesser degree among the SSRI-treated groups (SSRI users versus nonusers), uncontrolled confounding may have affected the risk estimates. Not adjusting for protective factors that were more likely to affect SSRI users would have tended to bias these results towards not finding an association (towards the null), whereas not adjusting for factors that increase breast cancer risk that were more common among the SSRI users would tend to bias the results towards overestimating an association between SSRI use and breast cancer.

In addition, the analysis did not account for the use of other medications associated with increased risk or decreased risk of breast cancer. If these medications were used to a greater or lesser degree among SSRI users, then risk estimates may have been affected by uncontrolled confounding.

As described previously, this lack of information on potential confounders will be addressed in stage two when detailed information related to established risk factors for breast cancer is being collected by telephone interviews on a sub-set of the stage one subjects. The final analysis of the two-stage study will consider these established factors as potential confounders and also include several medications that have been shown in previous experimental and epidemiologic studies to be associated with breast cancer.

(c) Confounding by indication: ‘Confounding by indication’, which is considered to be the most important confounding factor in pharmacoepidemiology (Csizmadi et al., 2005, p.799), is another source of potential uncontrolled confounding in this study. The most common cause of confounding by indication arises when the disease or underlying
condition for which the drug was prescribed acts as a confounder in the relationship of interest (Csizmadi et al., 2005, p.799). In this situation, to be a true confounder, the underlying condition must be associated with SSRI use, be an independent risk factor for breast cancer and not be on the causal pathway.

Since the SK prescription database and other sources of data for this thesis project did not contain information related to indications (mainly depression) for prescribing SSRIs, the analysis did not control for these conditions. Based on physician prescribing data (1997–1999) from the National Disease Therapeutic Index physician survey, the majority of prescriptions for SSRIs (fluoxetine, sertraline, paroxetine and citalopram) were prescribed for the treatment of depressive disorders (77.6%: range 74.0-80.3%) (Loosbrock et al., 2002). Epidemiologic studies looking at the association between depression and cancer have reported inconclusive results, with findings of marginally increased cancer risk associated with chronic depression (Jacobs and Bovasso, 2000; Penninx et al., 1998) and reports of no increased risk (Butow et al., 2000; Hahn and Petitti, 1988). A meta-analysis by McGee et al. (1994) based on prospective longitudinal studies reported a borderline statistically significant association (OR=1.14, 95% CI: 0.99-1.30) between depression diagnosed at the time of study enrolment and the development of cancer 10 or more years later. Further, there is evidence to suggest that some risk factors for breast cancer (heavy alcohol use and decreased physical activity) may be associated with increased rates of depression and therefore potentially linked with increased SSRI use (Wang & Patten, 2001; Dunn et al., 2001).

Given this evidence to suggest that SSRIs are mainly prescribed for the treatment of depression and that the risk of breast cancer may be increased with chronic depression, risk estimates for this study may have been confounded by depression, biasing the results.
towards finding an association between long-term SSRI use and breast cancer. However, since these results demonstrate no increased risk, this bias is of less concern.

Further, physician prescribing practices may be influenced by a patient’s clinical or behavioural characteristics; health-related factors that may also be independent predictors of the study outcome (Schneeweiss, 2006). Failing to control for such factors can be another source of uncontrolled confounding due to confounding by indication. Within the context of this study, if a physician chose whether or not to prescribe an SSRI based on unmeasured patient characteristics related to breast cancer development (including protective effects or risk factors), then results of this study may have been biased. However, based on the fact that a potential association between the use of SSRIs and breast cancer is not well known, it is unlikely that the presence of established risk factors for breast cancer such as smoking, family history and reproductive factors would influence physician prescribing practices. Further, among patients diagnosed with depression, factors shown to be associated with an increased likelihood of prescribing SSRIs as compared to non-SSRI antidepressants include white race (Melfi et al., 2000) younger age (Sleath & Shih, 2003; Olfson et al., 1998), non-psychiatrist (family physician) as the attending physician (Sleath & Shih, 2003), private insurance coverage (Sleath & Shih, 2003) and less severe depressive symptoms (Sleath & Shih, 2003; Olfson et al., 1998; Mojtabai, 2002). No evidence was found to suggest that the presence or absence of most risk factors for breast cancer (family and reproductive history, lifestyle factors) was associated with ‘selective’ prescribing practices (choosing to prescribe an SSRI based on these factors). However, regardless of whether or not there is the potential for an association between SSRI use and these confounders in the general population, it is imperative to assess these relationships in the study sample in order to interpret risk.
estimates (Breslow and Day, 1980). Therefore, information related to all factors that could potentially be related to SSRI use and breast cancer development is being collected by telephone interview from a sub-sample of stage one subjects for the larger two-stage study.

**Detection bias:** Detection bias, often referred to as differential follow-up, which occurs when identification of the outcome is influenced by the presence of the exposure, is a source of differential misclassification of the outcome in pharmacoepidemiologic studies (Csizmadi et al., 2005, p.797). In the context of this study population, for SSRI users, regular appointments to monitor SSRI use could potentially have led to an apparent excess number of diagnosed breast cancer cases in the SSRI user group. On the other hand, less frequent medical follow-up for nonusers of SSRIs may have led to undetected breast cancers in this group, resulting in potential misclassification of controls. Therefore, if SSRI users as compared to nonusers in this study population had more frequent medical care and this resulted in more chance of breast cancer diagnosis in the SSRI user group (differential follow-up), then risk estimates may have been falsely elevated.

**Limitations of exposure window approach:** The main exposure window of interest in this study included all exposures within the period of time more than two years prior to index date. Risk estimates associated with the two shorter exposure windows (2-7 and >7 years prior to index date) were intended to assess whether or not there was a suggested increased risk with heavy SSRI use specifically during each of these intervals, therefore reflecting a potential tumour promoter effect of SSRIs.
This time windows approach used to estimate risk associated with each of the two shorter exposure windows (2-7 and >7 years prior to index date) did not take into account use during the other of the two time periods and therefore resulted in the SSRI exposed group being a mix of individuals who have varying total SSRI exposures over the full exposure period more than two years prior to index date (Lefebvre et al., 2006). For instance, for the time period 2-7 years prior to index date, the SSRI user group would consist of a mix of new users (those who only used SSRIs in the 2-7 interval) and previous users consisting of those who in addition to using SSRIs in the 2-7 year interval, also used SSRIs during the period of time more than seven years prior to index date.

In order to take into account whether or not individuals had previous or subsequent SSRI use, the next step would be to assess risk during each exposure window for the two SSRI user groups separately (new users and previous users). An alternative approach, based on Miettinen’s idea (1985) that exposures during different time intervals represent separate determinants of risk, would consist of adjusting for confounding by including SSRI exposures during different exposure windows in the same statistical model as done in previous pharmacoepidemiologic studies (Sharpe et al, 2002; Tamim, 2003).

5.6 Summary explanation for null findings

In summary, these null findings may reflect the truth or chance. Or, if this study failed to detect a true increase in breast cancer risk associated with long-term SSRI use, these null findings may be explained by sources of information bias (nondifferential misclassification) and confounding. Also, since the prevalence of high SSRI exposure (≥36 prescriptions) was lower than anticipated (actual prevalence was 2.2% versus
anticipated 3% prevalence), it is expected that the power in this study was a bit less than 80%. In addition, for exposures to individual SSRIs, too few highly exposed cases meant that the analyses may have lacked power to detect small increases in risk.

Latency issues may also account for not observing an association. Even though the follow-up period for 99% of cases and controls was 15 or more consecutive years prior to index date, from a carcinogenic perspective, the latency period between SSRI exposure and breast cancer development may be such that not enough time had passed to assess risk associated with heavy use during the most meaningful exposure periods. In addition, to ensure a long-term duration of exposure category with an adequate number of subjects, the highest exposure group included individuals with 36 or more prescriptions. This total duration of use may have been too short to see a potential carcinogenic effect of long-term SSRI use.

5.7 Suggestions for future research

It is important to note that these results represent preliminary findings based on the first three years of subject accrual for the larger two-stage study. Many of the limitations of this thesis research will be addressed in future analysis of the complete two-stage study which will include two more years of data, for a total of five years of subject accrual. Confounder information related to other indicators of SES (education and income level), reproductive history, health care utilization (visits to family doctor and mammogram use), and indications for SSRI use is being collected through telephone interviews on a sub-sample of stage one subjects. Each of these factors, as well as the use of other medications that have suggested relationships with breast cancer, will be assessed as potential confounders in the final analysis. Also, analysis will be repeated for case
groups defined by histology and stage at diagnosis. Future analyses will potentially include a sufficient number of highly exposed cases and controls to allow an evaluation of risk associated with even longer durations of use and for heavier use overall and during specific exposure windows.

The relationship between breast cancer risk and SSRI use could be further explored using other study methods besides the traditional epidemiologic approaches using self-reports and databases to ascertain SSRI exposure information. As described previously, there is evidence to suggest that SSRIs may have the potential to enhance breast cell proliferation and growth indirectly through an SSRI-mediated increase in the hormone, prolactin. Biomarkers (biochemical measures that are used to detect cellular, biochemical or molecular alterations in biologic media or human tissues) reflect exposures that are more proximal to the disease occurrence, and therefore represent objective measures of exposure levels at intermediate stages of a causal mechanism. Within the context of SSRI exposures and breast cancer risk, measurement of prolactin levels may represent a valid and meaningful biomarker of the biological activity of ingested SSRI exposure. Several single case reports and other uncontrolled studies that assessed changes in prolactin levels with SSRI use have indicated that all SSRIs have the potential to cause an increase in circulating prolactin levels (Emiliano and Fudge, 2004). Therefore, further controlled studies designed to characterize changes in prolactin levels associated with long-term SSRI use may provide objective and valuable information towards understanding the underlying etiologic relationships between breast cancer and long-term SSRI use (Emiliano and Fudge, 2004).
5.8 Summary and Contribution of this Thesis Research

Faich and Stadel (1989) believe that “the future of pharmacoepidemiology is largely dependent on increased and careful use of automated databases”. This study incorporated a prescription database approach to ascertaining exposure to SSRIs and confounding information was collected on a limited set of potential confounders from the SK population and prescription databases. Results of this thesis project which are based on the first three years of subject accrual for the larger two-stage project do not provide evidence to suggest that the risk of breast cancer is increased with the long-term use of SSRIs combined or with the use of individual SSRI drugs. However, these preliminary results must be interpreted with caution as they are subject to the limitations described in this chapter. In addition, very few observational studies have investigated the long-term safety of SSRIs and therefore direct comparisons with previous research are difficult.

In summary, this thesis research has presented a framework for conceptualizing SSRI exposure for the two-stage study which will facilitate and drive the final analysis of the complete stage one data. There is an urgent need to assess this biologically plausible association because even a modest increase in breast cancer risk due to long-term SSRI use would have major public health implications given projections of increasing use in the female population for the treatment of many medical and psychiatric conditions.
References


LaBella FS, Brandes LJ. Enhancement of tumor growth by drugs with some common molecular actions. Molecular Carcinogenesis 1996;16(2):68-76.


Schlesselman JJ. Case-control studies; design, conduct, analysis. New York: Oxford University Press, 1982:144-152.


APPENDIX 1

Queen’s University Research Ethics Board Approval (& Renewals)
QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD

Queen's University, in accordance with the "Tri-Council Policy Statement, 1998" prepared by the Medical Research Council, Natural Sciences and Engineering Research Council of Canada and Social Sciences and Humanities Research Council of Canada requires that research projects involving human subjects be reviewed annually to determine their acceptability on ethical grounds.

A Research Ethics Board composed of:

Dr. A.F. Clark  Emeritus Professor, Department of Biochemistry, Faculty of Health Sciences, Queen's University (Chair)
Dr. S. Burke  Emeritus Professor, School of Nursing, Queen's University
Rev. T. Deline  Community Member
Dr. M. Evans  Community Member
Dr. M. Green  Assistant Professor, Department of Family Medicine, Queen's University
Ms. T.C. Knott  Research & Evaluation, Southeastern Regional Geriatric Program, Providence Continuing Care Centre – St. Mary's of the Lake Hospital Site
Dr. J. Low  Emeritus Professor, Department of Obstetrics and Gynaecology, Queen's University and Kingston General Hospital
Dr. H. Murray  Assistant Professor, Department of Emergency Medicine, Queen's University
Dr. W. Racz  Emeritus Professor, Department of Pharmacology & Toxicology, Queen's University
Dr. H. Richardson  Assistant Professor, Department of Community Health & Epidemiology Project Coordinator, NCIC CTG, Queen's University
Dr. B. Simchison  Assistant Professor, Department of Anesthesiology, Queen's University
Dr. A.N. Singh  WHO Professor in Psychosomatic Medicine and Psychopharmacology Professor of Psychiatry and Pharmacology Chair and Head, Division of Psychopharmacology, Queen's University
Dr. S. Taylor  Director, Office of Bioethics, Queen's University and Kingston General Hospital; Associate Professor, Department of Medicine, Queen's University
Ms. K. Weisbaum  LL.B. and Adjunct Instructor, Department of Family Medicine (Bioethics)

has examined the protocol for the project entitled "Selective serotonin reuptake inhibitors and breast cancer: a record linkage study" as proposed by Ms. Janet Ashbury, Dr. Kristan Aronson and Dr. Will King of the Department of Community Health and Epidemiology at Queen's University and considers it to be ethically acceptable. This approval is valid for one year. If there are any amendments or changes to the protocol affecting the subjects in this study, it is the responsibility of the principal investigator to notify the Research Ethics Board. Any unexpected serious adverse event occurring locally must be reported within 2 working days or earlier if required by the study sponsor. All other serious adverse events must be reported within 15 days after becoming aware of the information."

Chair, Research Ethics Board  Date

Examiner's signature  Oct 28, 2005

Original to Investigator - Copy to Department Head - Copy to Hospital(s) - P&T - File Copy

EPID-209-05
EX
QUEEN'S UNIVERSITY HEALTH SCIENCES AND AFFILIATED TEACHING HOSPITALS
ANNUAL RENEWAL

Queen's University, in accordance with the "Tri-Council Policy Statement, 1998" prepared by the Medical Research Council, Natural Sciences and Engineering Research Council of Canada and Social Sciences and Humanities Research Council of Canada requires that research projects involving human subjects be reviewed annually to determine their acceptability on ethical grounds.

A Research Ethics Board composed of:

Dr. A.F. Clark Emeritus Professor, Department of Biochemistry, Faculty of Health Sciences, Queen's University (Chair)

Dr. S. Burke Emeritus Professor, School of Nursing, Queen's University

Rev. T. Deline Community Member

Dr. M. Evans Community Member

Mr. C. Kenny Community Member

Dr. J. Low Professor, Department of Obstetrics and Gynaecology, Queen's University and Kingston General Hospital

Dr. H. Murray Assistant Professor, Department of Emergency Medicine, Queen's University

Dr. W. Racz Emeritus Professor, Department of Pharmacology & Toxicology, Queen's University

Dr. H. Richardson Assistant Professor, Department of Community Health & Epidemiology Project Coordinator, NCIC CTG, Queen's University

Dr. B. Simchison Assistant Professor, Department of Anaesthesiology, Queen's University

Dr. A.N. Singh WHO Professor in Psychosomatic Medicine and Psychopharmacology Professor of Psychiatry and Pharmacology Chair and Head, Division of Psychopharmacology, Queen's University Director & Chief of Psychiatry, Academic Unit, Quinte Health Care, Belleville General Hospital

Dr. M. Sommerfeld Physician and Assistant Professor, Department of Family Medicine, Queen's University

Ms. K. Weisbaum LL.B. and Adjunct Instructor, Department of Family Medicine (Bioethics)

has reviewed the request for renewal of Research Ethics Board Approval for the project "Selective serotonin reuptake inhibitors and breast cancer: a record linkage study," as proposed by Dr. J. Ashbury, Dr. K. Armonon and Dr. W. King of the Department of Community Health and Epidemiology at Queen's University. The approval is renewed for one year, effective October 28, 2006. If there are any further amendments or changes to the protocol affecting the subjects in this study it is the responsibility of the principal investigator to notify the Research Ethics Board. Any unexpected serious adverse event occurring locally must be reported within 2 working days, or earlier if required by the study sponsor. All other adverse events must be reported within 15 days after becoming aware of the information.

Albert Clark
Chair, Research Ethics Board

Aug 10, 2006

Date

REB# EPID-209-05
QUEEN'S UNIVERSITY HEALTH SCIENCES AND AFFILIATED TEACHING HOSPITALS
ANNUAL RENEWAL
Queen's University, in accordance with the "Tri-Council Policy Statement, 1998" prepared by the Medical Research Council, Natural Sciences and Engineering Research Council of Canada and Social Sciences and Humanities Research Council of Canada requires that research projects involving human subjects be reviewed annually to determine their acceptability on ethical grounds.

A Research Ethics Board composed of:

Dr. A.F. Clark  Emeritus Professor, Department of Biochemistry, Faculty of Health Sciences, Queen's University (Chair)

Dr. S. Burke  Emeritus Professor, School of Nursing, Queen's University

Rev. T. Beline  Community Member

Dr. M. Evans  Community Member

Mr. C. Kenny  Community Member

Ms. T.C. Knott  Research & Evaluation, Southeastern Regional Geriatric Program, Providence Continuing Care Centre - St. Mary's of the Lake Hospital Site

Dr. J. Low  Emeritus Professor, Department of Obstetrics and Gynaecology, Queen's University and Kingston General Hospital

Dr. W. Racz  Emeritus Professor, Department of Pharmacology & Toxicology, Queen's University

Dr. H. Richardson  Assistant Professor, Department of Community Health & Epidemiology Project Coordinator, NCIC CTG, Queen's University

Dr. B. Simchison  Assistant Professor, Department of Anaesthesiology, Queen's University

Dr. A.N. Singh  WHO Professor in Psychosomatic Medicine and Psychopharmacology Professor of Psychiatry and Pharmacology Chair and Head, Division of Psychopharmacology, Queen's University Director & Chief of Psychiatry, Academic Unit, Quinte Health Care, Belleville General Hospital

Dr. E. Tsai  Assistant Professor, Department of Paediatrics and Office of Bioethics, Queen's University

Ms. K. Weisbaum  L.L.B. and Adjunct Instructor, Department of Family Medicine (Bioethics)

has reviewed the request for renewal of Research Ethics Board approval for the project "Selective Serotonin Reuptake Inhibitors and Breast Cancer: A Record Linkage Study" as proposed by Ms. Janet Ashbury, Dr. K. Aronson and Dr. W. King of the Department of Community Health and Epidemiology, at Queen's University. The approval is renewed for one year, effective October 28, 2007. If there are any further amendments or changes to the protocol affecting the subjects in this study, it is the responsibility of the principal investigator to notify the Research Ethics Board. Any unexpected serious adverse event occurring locally must be reported within 2 working days or earlier if required by the study sponsor. All other adverse events must be reported within 15 days after becoming aware of the information.

Chair, Research Ethics Board  Aug 15, 2007

ORIGINAL TO INVESTIGATOR - COPY TO DEPARTMENT HEAD - COPY TO HOSPITAL(S) - FILE COPY
Renewal 1[ ] Renewal 2 [ x ] Extension [ ]

REB# EPID-209-05

113
APPENDIX 2

Detailed description of case-control data linkage sampling method

This section will describe in more detail the complex method of case and control selection. In order to ascertain case and control eligibility according to age, prior cancer diagnosis and prescription coverage criteria, cases and controls were selected for this thesis project using the following steps. (See Figure 10)

Case selection process

1. SCA Breast cancer file: The SCA identified 1,935 breast cancer cases meeting the following criteria:
   - ICDO code = C50 (breast cancer)
   - Behaviour = 3 (malignant)
   - Diagnosis date from 2003 to 2005
   - Subject residing in SK at time of diagnosis
   - Death certificate only cases excluded

   The SCA then removed cases aged <26 years or >81 years of age at the time of diagnosis (n=244 cases), leaving 1,691 cases. (It was necessary to widen the case age range at this step since SK Health matches date of birth within plus or minus two years.) The SCA then removed breast cancer cases with a prior cancer diagnosis within the previous 10 years (except for non-melanoma skin cancer and cervical cancer in situ). The final SCA breast cancer file included 1,533 breast cancer cases corresponding to 1,496 individuals. (There were 36 individuals with more than one breast cancer diagnosis.
recorded in this file: 35 people had two records and one person had three records with the same index date for a total of 37 (35+2) additional breast cancer diagnoses recorded.

2. Eligible case file: The SCA breast cancer file was forwarded to SK Health which confirmed health services number (HSN) and date of birth (DOB) for breast cancer cases diagnosed between January 1, 2003 and December 31, 2005. Individuals with registered Indian status were removed from the file (not eligible for SK Health coverage and therefore not included in the prescription database). Cases with fewer than 10 consecutive years of SK Health coverage prior to index date were also excluded from the file. The eligible case file had 1,305 separate breast cancer diagnoses recorded corresponding to 1,273 individuals. There were 31 women with more than one breast cancer diagnosis recorded in this file: 30 women had two breast cancer diagnoses and one woman had three breast cancer diagnoses recorded on the same day for a total of 32 extra (30+2) breast cancer diagnoses recorded. These individuals with more than one breast cancer diagnoses represented women diagnosed on the same day with 2 (or 3) separate tumours in separate quadrants of the breast. Therefore, the eligible case file included all women diagnosed with primary invasive breast cancer between January 1, 2003 and December 31, 2005, who were aged 28-79, with no cancer diagnosis in previous 10 years and eligible for prescription coverage for at least 10 consecutive years.

Control selection process

1. Potential control pool file: Initially, all SK women who were born between January 1, 1923 and December 31, 1979 and had continuous SK Health coverage starting on or before December 31, 1997 were identified from the SK population registry. Individuals with registered Indian status were removed from the file as well as those whose coverage
ended before January 1, 2003. Those remaining are referred to as **potential control pool file** (N=262,697).

2. **Prior cancer file**: The SCA matched this potential control pool file to the cancer registry in order to identify those controls with any cancer diagnosed from January 1, 1993 to December 31, 2005 (excluding non-melanoma skin cancer or cervical carcinoma *in situ*). This file is referred to as the **prior cancer file** (N=14,283).

**Case-control selection process for each 6-month accrual period**

Using the first 6-month period of time (January 1, 2003 – June 30, 2003) as an example, this section will describe in detail the case and control selection process for each 6-month accrual period. Three files that are described above were used in this process: the **eligible case file** (cases diagnosed between January 1, 2003 and December 31, 2005), the **potential control pool file** (all women born between 1923 and 1979 with continuous coverage as described above) and the **prior cancer file** (all women identified from the potential control pool file with any cancer diagnosed between January 1, 1993 and December 31, 2005).

**Case/Control selection for first six month time period (Jan 1, 2003 – June 30, 2003)**

A: From the **eligible case file**, all cases diagnosed between January 1, 2003 and June 30, 2003 were selected.

B: Sub-set of controls from **prior cancer file**: with a cancer diagnosis date between January 1, 1993 and June 30, 2003 were selected.

C: Subjects from steps A and B were merged. This **merged cancer file** included all breast cancer cases diagnosed between January 1, 2003 and June 30, 2003 and all women
with a prior cancer diagnosis between January 1, 1993 and June 30, 2003. This step was done in order to confirm that all subjects in A were included in B (that is: to be sure that no breast cancer cases were missing from A).

D: A sub-set of controls was then selected from the potential control pool file which included all controls who were at least 28 years of age as of January 1, 2003 and who were less than or equal to 79 years of age as of June 30, 2003 with at least 10 years of continuous coverage prior to January 1, 2003 and SK Health coverage during the entire 6-month period of January – June 2003.

E: The merged cancer file produced in step C (by merging A and B) was combined with the sub-set of controls selected in step D and matching records were deleted, creating the potential eligible control group file. Effectively, this removed all potential controls with a cancer history from January 1, 1993 to June 2003 leaving those controls who were eligible for control selection for the first 6-month period.

F: Using the potential eligible control group file from step E, a randomly assigned index date between January and June 2003 was assigned to each of the control subjects and age was calculated as of that index date. From this file, the required number of controls was randomly sampled, 10 controls per case, frequency matched in 5-year age groups to the actual case age distribution for the cases with index dates during the same 6-month time period.

To select controls for the five subsequent 6-month intervals, the process followed the same steps, whereby 10 controls per case were selected, frequency matched to the cases by age in 5-year age groups. Each selection process started with the original potential control pool file, effectively placing controls back into the ‘pool’ for each 6-
month accrual process. Therefore, controls could be selected as a control again or become a case for subsequent 6-month accrual periods.

**Summary:** Using a series of data linkage steps, eligible cases and controls were selected in 6-month groupings over the 3-year case accrual period. The final case-control sample for this thesis project consisted of 1,273 cases and 12,730 controls.
Figure 10: Case-control sampling method

**All breast cancer cases:** diagnosed in 2003-2005 identified from SCA cancer registry (N=1,935)

**SCA Breast Cancer File:** All primary invasive breast cancer cases; diagnosed Jan/03 – Dec/05, ages 26-81 (SK Health matches DOB to within 2 yrs), no cancer diagnosis in previous 10 years (N=1,496)

**Eligible Case File:** Confirm HSN, DOB, (within 2 years), coverage for $\geq$ 10 yrs, Breast cancer diagnosed Jan/03 – Dec/05 (N=1273)

**Potential Control Pool File** (SK Health), all women born between Jan/23- Dec/79, pool for controls for all parts of study, continuous SK health coverage starting on or before December 31, 1997 (N=262,697)

**Prior Cancer File:** (SCA) Identifies all women in potential control file with any cancer diagnosed between Jan/93 and Dec/05 (N=14,283 records)

**First 6-month sampling procedure:** (Jan/03-June/03)

**A: Breast cancer cases** diagnosed Jan/03 – June/03 (age 28-79, coverage $\geq$ 10 yrs & no prior CA)

**Combine A & B** To confirm that all cases included

**C: Merged cancer file:** Includes all breast cancer cases (Jan/03-June/03) AND women with prior cancer history (Jan/93-June/03)

**Combine C and D: delete matches**

**D: Sub-set of controls from potential control pool:** women age 28 by Jan/03: & $\leq$ 79 by Jun/03, Coverage for $\geq$ 10 yrs prior to Jan/03 Health coverage during entire 6-month period

**E: Potential eligible control file:** all with coverage $\geq$ 10 yrs & no prior CA (Jan/93-June/03), randomly assign index date: Jan–June ’03

**F: Controls:** Randomly select 10 controls per case without replacement from potential eligible control file (index date: Jan/03 – June/03)
APPENDIX 3: Prevalence Data Tables

Table 3.1: Prevalence of SSRI use by year among controls (Exact data for Figure 7)

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Prevalence (%) Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>0.2</td>
</tr>
<tr>
<td>1990</td>
<td>0.9</td>
</tr>
<tr>
<td>1991</td>
<td>1.2</td>
</tr>
<tr>
<td>1992</td>
<td>1.5</td>
</tr>
<tr>
<td>1993</td>
<td>1.6</td>
</tr>
<tr>
<td>1994</td>
<td>2.8</td>
</tr>
<tr>
<td>1995</td>
<td>3.1</td>
</tr>
<tr>
<td>1996</td>
<td>3.7</td>
</tr>
<tr>
<td>1997</td>
<td>4.0</td>
</tr>
<tr>
<td>1998</td>
<td>4.2</td>
</tr>
<tr>
<td>1999</td>
<td>4.8</td>
</tr>
<tr>
<td>2000</td>
<td>5.5</td>
</tr>
<tr>
<td>2001</td>
<td>6.0</td>
</tr>
<tr>
<td>2002</td>
<td>6.5</td>
</tr>
<tr>
<td>2003</td>
<td>6.8</td>
</tr>
<tr>
<td>2004</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Table 3.2: Prevalence of individual SSRI use by year among SSRI users (controls) (Exact data for Figure 8)

<table>
<thead>
<tr>
<th>YEAR</th>
<th>Fluoxetine (%) Controls</th>
<th>Fluvoxamine (%) Controls</th>
<th>Sertraline (%) Controls</th>
<th>Paroxetine (%) Controls</th>
<th>Citalopram (%) Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>100</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>73.1</td>
<td>34.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>66.7</td>
<td>33.8</td>
<td>6.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>55.3</td>
<td>20.5</td>
<td>11.2</td>
<td>22.5</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>49.2</td>
<td>17.8</td>
<td>14.1</td>
<td>26.1</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>47.4</td>
<td>17.5</td>
<td>12.2</td>
<td>27.4</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>41.8</td>
<td>18.2</td>
<td>17.5</td>
<td>29.8</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>37.4</td>
<td>14.7</td>
<td>21.1</td>
<td>32.5</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>32.1</td>
<td>12.4</td>
<td>24.5</td>
<td>37.2</td>
<td>1.1</td>
</tr>
<tr>
<td>2000</td>
<td>24.6</td>
<td>9.9</td>
<td>22.7</td>
<td>40.7</td>
<td>10.3</td>
</tr>
<tr>
<td>2001</td>
<td>21.3</td>
<td>6.6</td>
<td>22.4</td>
<td>40.3</td>
<td>15.3</td>
</tr>
<tr>
<td>2002</td>
<td>20.4</td>
<td>6.4</td>
<td>18.0</td>
<td>40.2</td>
<td>19.9</td>
</tr>
<tr>
<td>2003</td>
<td>19.4</td>
<td>6.3</td>
<td>17.7</td>
<td>38.4</td>
<td>25.2</td>
</tr>
<tr>
<td>2004</td>
<td>18.6</td>
<td>4.6</td>
<td>14.4</td>
<td>36.3</td>
<td>29.4</td>
</tr>
</tbody>
</table>
**APPENDIX 4: Alternative Confounder Assessment**

Table 4.1: Assessment of potential confounders using the change-in-estimate method

<table>
<thead>
<tr>
<th>Covariate</th>
<th>OR without factor (%Δ)</th>
<th>OR without factor (%Δ)</th>
<th>OR without factor (%Δ)</th>
<th>OR without factor (%Δ)</th>
<th>OR without factor (%Δ)</th>
<th>OR without factor (%Δ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>1.077 (1.01%)</td>
<td>1.076 (1.01%)</td>
<td>1.080 (0.92%)</td>
<td>1.080 (0.64%)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Income support status</td>
<td>1.082 (0.55%)</td>
<td>1.081 (0.55%)</td>
<td>1.078 (0.83%)</td>
<td>1.078 (0.83%)</td>
<td>1.062 (1.67%)</td>
<td>X</td>
</tr>
<tr>
<td>Residence status</td>
<td>1.091 (0.28%)</td>
<td>1.090 (0.28%)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>OC use</td>
<td>1.087 (0.09%)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>HT use</td>
<td>1.111 (2.11%)</td>
<td>1.108 (1.93%)</td>
<td>1.113 (2.10%)</td>
<td>1.109 (2.02%)</td>
<td>1.102 (2.04%)</td>
<td>1.083 (1.98%)</td>
</tr>
<tr>
<td>Age</td>
<td>1.077 (1.01%)</td>
<td>1.083 (0.37%)</td>
<td>1.087 (0.28%)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>OR for each model</td>
<td>1.088</td>
<td>1.087</td>
<td>1.090</td>
<td>1.087</td>
<td>1.080</td>
<td>1.062</td>
</tr>
<tr>
<td>Total %Δ in estimate accrued from the start of the process (model with all confounders controlled)</td>
<td>Fully adjusted model</td>
<td>0.09%</td>
<td>0.18%</td>
<td>0.09%</td>
<td>0.74%</td>
<td>2.39%</td>
</tr>
</tbody>
</table>