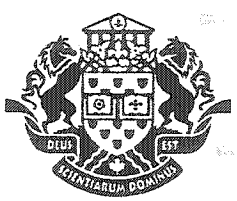


Université d'Ottawa · University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

Alexander TSERTSVADZE

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

M. Sc. (Epidemiology)

GRADE - DEGREE

Department of Epidemiology

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Recreational physical activity, relative body weight and risk of bladder cancer : a case-control study

C. Nair

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

N. Birkett and I. McDowell

CO-DIRECTEUR DE LA THÈSE - THESIS CO-SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

G. Wells

D. Wigle

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

RECREATIONAL PHYSICAL ACTIVITY, RELATIVE BODY WEIGHT AND RISK OF BLADDER CANCER: A CASE-CONTROL STUDY

BY ALEXANDER TSERTSVADZE

THESIS SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE MSc. DEGREE IN EPIDEMIOLOGY

UNIVERSITY OF OTTAWA

APRIL 2004

Your file Votre référence
ISBN: 0-494-01622-1
Our file Notre référence
ISBN: 0-494-01622-1

ABSTRACT

Bladder cancer is the second most frequently diagnosed neoplasm of the urinary tract. Effects of physical activity and body fat mass on risk of bladder cancer are not well studied. Physical activity and/or body weight may exert their effects on bladder cancer risk through hormonal, immune, or yet other unknown causal pathways. This thesis, using a case-control study design, investigated the association between recreational physical activity, relative body weight (BMI), and risk of bladder cancer in men and women separately. The analysis was based on a portion of the National Enhanced Cancer Surveillance System dataset. The self-reported data had been collected from bladder cancer cases and population controls during 1994-1997 in seven Canadian Provinces. The analyzed samples were 2,312 males (670 cases and 1,642 controls) and 1,824 women (359 cases and 1,465 controls). In men and women, the adjusted ORs (highest vs. lowest quartile) for physical activity were 1.24 (95% CI: 0.93, 1.66) and 0.76 (95% CI: 0.52, 1.21), respectively. The corresponding ORs for BMI were 1.29 (95% CI: 0.96, 1.72) and 1.07 (95% CI: 0.74, 1.55), respectively. This study found that parous women were at reduced risk of bladder cancer compared to nulliparous women (OR = 0.56, 95% confidence interval: 0.39, 0.80). In agreement with other studies, certain occupations and cigarette smoking were associated with increased risk of bladder cancer. In men, but not in women, coffee consumption was associated with a slightly increased risk of bladder cancer (≥ 1 cup per day vs. never: OR=1.30, 95% CI: 1.00, 1.68). There was not enough evidence in this study to conclude that recreational physical activity and BMI were related to risk of bladder cancer.

ACKNOWLEDGEMENTS

Without contribution of many individuals, it would not have been possible to complete this thesis.

I find it hard to express my deep gratitude towards the thesis supervisors Dr. Rama Nair, Dr. Anne-Marie Ugnat, Dr. Nicholas Birkett, and Dr. Ian McDowell. Many thanks for your help and expert guidance and for making this project come to life. Each of you made an invaluable and unique contribution to the thesis.

Dr. Rama Nair, I will not forget your continuous genuine human to human support throughout my studies at the department. Your presence made my and many other students' hectic and stressful lives a more pleasant experience. Many thanks for providing me with your invaluable knowledge and experience.

Dr. Anne-Marie Ugnat, you are a very kind soul and I am deeply grateful for everything what you have done for me throughout my studies at the department. I also thank you for introducing me to Health Canada without which the implementation of this project would not have been possible. I am extremely grateful for your continuous guidance and expertise not only with the thesis but also with aspects of my career and well-being.

Dr. Nicholas Birkett, please accept my deepest appreciations for your invaluable expertise, time, and patience. During all this time, you were the driving force of my scientific inspiration. It is hard to describe how much I learned from you. This thesis would not be as good as it is now without your knowledge, cautiousness, and meticulous or detailed view of data. Thank you again for your interest in this project.

Dr. Ian McDowell, many thanks for being one of my supervisors. I am extremely thankful for your help, knowledge, and unique expertise to make this thesis better. I am deeply grateful for your continuous encouragement and interest in my career. The student life at the department would not have been the same fun without your humor and friendliness.

Many thanks to Dr. George Wells, Dr. Yue Chen, and Dr. Robert Spasoff for their overall support throughout my studies at the department.

Dear Fay Draper, you are the kindest soul. Words cannot describe my appreciation for your presence.

Dear Sylvie Desrochers, Anne Grey, and Mariella Peca, thank you all very much, through my studies and long hours at the department, you made me feel at home.

My friends and colleagues, Amira Ali, Supriya Mayank, Kevin Brand, and Ken Johnson, thank you for your suggestions and overall support.

Thank you all at the Thomas Chalmers Centre for Systematic Reviews

To my wife Tea and my daughter Nina, your tolerance and support throughout my courses and thesis have given me the will to do my best. Thank you for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

1.1. The Epidemiology of Bladder Cancer

Bladder cancer is the second most frequently diagnosed neoplasm of the urinary tract. In 1999, nearly 263,000 cases of bladder cancer were diagnosed worldwide [1-2]. The highest incidence rates of this type of malignancy are observed in Western Europe and North America [2-4]. The occurrence of the disease amongst whites is twice as frequent as that amongst blacks [5]. Bladder cancer is the fourth most frequently diagnosed malignancy in Canadian men. An estimated 5,000 new cases (3,700 male and 1,300 female) of bladder cancer were expected to occur in Canada in 2003. The annual age-standardized incidence rate (standardized to the age distribution of the 1991 Canadian population) of bladder cancer in males for the year 2003 was 23 per 100,000 Canadian men. The corresponding annual rate for females was 6 per 100,000 Canadian women [6]. According to Patton et al., bladder cancer was the fourth leading cause of cancer deaths in men in the United States.

The risk of bladder cancer increases sharply with age. More than 65% of bladder cancer cases are persons older than 65 years [1]. Men, compared with women, are approximately two-to-four times more likely to develop bladder cancer, rendering sex as an important risk factor [2; 5-7]. The excess risk in males could partially be explained by their higher prevalence of tobacco smoking, more frequent employment in high-risk occupations, and more intensive exposures to occupational carcinogens. However, two studies found that even after the adjustment for these factors, the excess risk associated with being male persisted [8-9]. Based on this and other findings from human and animal studies [10-22], it was suggested that hormonal (sex-hormones) factors might play an additional and

independent role in the sex-specific rate differences and the etiology of bladder cancer, in general [8-10].

The well-established risk factors for bladder cancer are tobacco smoking, certain occupations (dye workers, truck drivers, hairdressers, petroleum workers, miners, textile workers, welders, and others), occupational exposures to certain chemicals (aromatic amines, rubber, painting, leather, aluminum), pharmaceutical drugs (Phenacetin and Cyclophosphamide), and urinary tract diseases (*Schistosoma Haematobium*) [2; 5; 7]. The epidemiologic evidence is still inconclusive regarding the etiologic roles of such factors as coffee consumption, water chlorination by-products, artificial sweeteners, diet, fluid intake, alcohol consumption, and family history in the carcinogenesis of bladder cancer [3; 5; 7; 23-26].

Ninety-three percent of bladder cancers are transitional cell carcinomas. Squamous cell carcinomas and adenocarcinomas account for approximately 3% of all bladder cancers [5; 23]. The majority of the neoplasms (about 80%) at first presentation are superficial and papillary. The remaining 20% of the tumours are muscle-invasive and are characterized with a worse prognosis [1].

1.2. Physical Activity, Body Weight, and Risk of Bladder Cancer

Physical activity and body weight are known risk factors for several human cancers. High levels of physical activity and lower weight have consistently been shown to be associated with a reduced risk of several cancers: colon, breast, endometrial, kidney, and

oesophagus [27-29]. Although, the exact biologic pathway between these factors and the carcinogenesis of cancer is not clear, it has been hypothesized that the effect of physical inactivity and excess body weight may be related to increased secretion of endogenous sex hormones (estrogens and androgens), plasma insulin, and insulin-like growth factor (IGF-I). Chronic exposure to increased levels of the above-mentioned factors has been implicated in carcinogenesis [27-31]. Physical inactivity has also been shown to be associated with decreased immune function, manifesting itself by reduced numbers of circulating natural killer cells (lymphocytes, granulocytes, monocytes, and macrophages). Reduced immune system capacity has been linked to carcinogenesis [32].

Several studies have demonstrated the association between hormonal factors and bladder cancer risk. Besides those studies [8-22] that have shown a link between the endogenous sex hormones and risk of bladder cancer, two other studies have found positive relationship between levels of insulin-like growth factor and risk of bladder cancer [33-34].

Given the findings that physical inactivity and excess body weight are involved in the metabolic alterations of endogenous hormones (sex hormones, plasma insulin, and insulin-like growth factor), which in turn are linked to risk of bladder cancer, it would be plausible that physical activity and body weight might exert their effects on bladder cancer risk through these hormonal or immunological pathways.

Only a few epidemiologic studies have investigated the associations of physical activity and/or body weight in relation to risk of bladder cancer [22; 35-41]. None of these studies

focused specifically on bladder cancer as the main outcome variable [42]. The results from these studies regarding the associations between levels of physical activity, body weight, and risk of bladder cancer are conflicting. Inadequate control for confounding and small numbers of bladder cancer cases could be responsible for the observed discrepancies across these studies.

Based on all the above-mentioned issues, epidemiologic investigation focusing exclusively on the potential effects of physical activity and body weight on risk of bladder risk is warranted. Moreover, the exploration of sex as an effect modifier of the association between physical activity/body weight and bladder cancer risk is needed. More research should be focused on the elucidation of the potential aetiological roles of hormonal factors (directly or through proxy measurements) in the development of human bladder cancer.

2. Literature Review

2.1. Biology and Clinical Features of Bladder cancer

2.1.1. Histopathology

Based on the World Health Organization (WHO) recommendations, epithelial bladder tumours are divided into four main histological types: transitional cell carcinoma, squamous cell carcinoma, adenocarcinoma, and undifferentiated carcinoma [43]. In North America and many parts of Western world, at least 90% of all bladder cancer cases are transitional cell carcinomas [3; 5; 44]. Squamous cell carcinomas account for approximately 2-7% [24; 44]. Only about 1% of all bladder cancers are comprised of adenocarcinomas [3; 5; 24; 44]. In Egypt, unlike North America and the Western world, squamous cell carcinomas account for about 80% of all malignant bladder tumours [24]. The majority (about 60%) of the transitional cell carcinomas are located in the mucosal layer or the lamina propria. About one third of all transitional cell carcinomas are multifocal [44].

2.1.2. Pathophysiology

Based on the mode of presentation, bladder tumours are usually divided into two broad groups: superficial and muscle-invasive [24; 44-45]. Superficial tumours are mostly papillary shaped and make up about 75-80% of all human bladder tumours. From twenty to thirty-five percent of all bladder tumours are muscle-invasive [1; 3; 24; 44]. Although superficial papillary bladder tumours are relatively benign (mostly non-invasive), with lower grade and stage at presentation, 10-15% of them develop invasion or metastases on average 15 years after the diagnosis. Muscle-invasive bladder tumours are more aggressive and are characterized with a poor grade of histological differentiation and advanced stage at the

first presentation. The majority of the patients with invasive bladder cancer have muscle invasion at the first presentation [24; 44]. It is not clear whether superficial and muscle-invasive groups of bladder cancer have similar or different etiology [44].

2.1.3. Signs, Symptoms, and Prognostic Factors

The first and most frequent sign of early bladder cancer is the presence of asymptomatic macroscopic hematuria or irritating voiding symptom. The proportion of bladder cancer patients presenting with vesical irritation may range from 30% to 60%. Other signs of bladder cancer include dysuria and incontinence. The patients with advanced bladder cancer may present with pelvic or bone pain. Tumour stage and grade are the most important prognostic factors for bladder cancer [24; 46].

2.1.4. Sex-Based Differences

Sex-specific differences in clinical, pathological characteristics, and natural history of the disease could provide important clues to the etiology of bladder cancer. Furthermore, the above-mentioned factors may yield additional knowledge or insight in the explanation of the observed sex-dependent variation in the risk of bladder cancer. A number of studies, using either newly diagnosed patients or cases obtained from cancer registries, have investigated distributions of clinical and pathological characteristics in males and females [46-49]. The results of these studies are conflicting. Two studies [46; 48] concluded that males had more advanced forms of the disease relative to females as demonstrated by a higher incidence of high-grade and high-stage tumours in men. In contrast, the other two studies [47; 49] suggested that male patients were diagnosed with lower staged disease than female

patients. For example, one study [49] found that the odds of being diagnosed with superficial bladder cancer were lower in females as compared to males (OR = 0.24, 95% CI: 0.06, 0.94). The other study [47] observed that women were more likely to be diagnosed at later stage and with more advanced disease than men (33.1% and 27.7% for women and men, respectively). It has been suggested that the higher proportion of advanced forms of disease amongst women could be due to the lower incidence of the disease in women; doctors may be less likely to consider the diagnosis of bladder cancer at their first visit. An alternative hypothesis considers the operation of sex-related biologic factors that could be responsible for the observed differences [47].

One study compared the frequency of commonly encountered symptoms (vesical irritation, hematuria, and dysuria) during the diagnosis of bladder cancer in men and women. The results of the study indicated that most of the symptoms did not differ across sex. However, women reported incontinence at first presentation more frequently than men [46].

2.2. Trends in Bladder Cancer Incidence and Mortality

In 1999, nearly 263,000 new cases of bladder cancer were diagnosed worldwide [1-2]. The highest incidence rates of bladder cancer are observed in Western Europe and North America. For example, amongst males, the highest age-adjusted (world standard) rates (1978 to1982) were recorded in Italy (26 per 100,000 person-years), Denmark (24 per 100,000 person-years), France (21 per 100,000 person-years), the U.S. (20 per 100,000 person-years), and Canada (20 per 100,000 person-years) [5].

North Africa and Western Asia are also high-risk areas [2-4]. For example, Egypt is an endemic area for *schistosoma haematobium*, which predisposes patients to bladder cancer [5].

Lower sex-specific age-adjusted incidence rates (ranging from 2 to 12 per 100,000 person-years) of bladder cancer are observed in Eastern Europe (Poland and Hungary) and some parts of Asia (Philippines, India, and Kuwait).

The geographic variation in bladder cancer incidence could partially be due to the differences in local practices of tumour registration. For example, registration of benign tumours or papillomas as cancer may vary according to a geographic area. Furthermore, for developing countries, under-coverage of cancer cases by local tumour registries could be a common problem [5].

Bladder cancer is the fourth most frequently diagnosed malignancy in Canadian men. An estimated 5,000 new cases (3,700 male and 1,300 female) of bladder cancer were expected to occur in Canada in 2003. The annual age-standardized incidence rate (standardized to the age distribution of the 1991 Canadian population) of bladder cancer in males for the year 2003 was 23 per 100,000 Canadian men. The corresponding annual rate for females was 6 per 100,000 Canadian women [6].

The estimated age-standardized bladder cancer mortality rates for 2003, in men and women, were 7 and 2 per 100,000 Canadians, respectively. It was estimated that in 2003,

1,100 men and 460 women would die from bladder cancer in Canada. According to Canadian statistics, between 1974 and 1998, the annual age-standardized incidence rates of male bladder cancer decreased from 26.4 to 23.5 per 100,000 Canadian men. Over the same period, the age-standardized mortality rates for males due to bladder cancer decreased from 8.5 to 7.5 per 100,000 Canadian men [6]. In the United States between 1969 and 1986, the annual age-standardized incidence rates of bladder cancer exhibited increasing trends in men and women of both race (whites and blacks). The large part of this rise in incidence appears to be a result of the increased detection of localized bladder cancer. In contrast, the corresponding mortality rates decreased over the same period [5].

2.3. Risk Factors

The well-established risk factors for bladder cancer are age, sex, tobacco smoking, occupational/chemical exposures, and certain urinary tract diseases. This section will briefly review the evidence for these factors. It will be followed by a review of several suspected risk factors for which the evidence is less convincing. A detailed review of the factors (physical activity, body weight, and hormonal factors) of primary interest to this thesis will be deferred to sections 2.7. and 2.8.

2.3.1. Sex, Race, and Age

Men are about two-to-four times more likely to get bladder cancer than women [2; 5-7]. Sex-specific differences with respect to tobacco smoking and employment in the high risk occupations explain only a part of the male excess risk for bladder cancer. It has been

suggested that sex-specific factors (i.e., endogenous hormones) might additionally explain these differences [8-22].

The incidence rate of bladder cancer in blacks is about half of that amongst whites [5]. Hispanic groups and American Indians have lower risks of bladder cancer than blacks [5]. The black-white difference in incidence rates persists after adjustment for social class, smoking, and occupation [5; 9].

For both, men and women, the risk of bladder cancer increases sharply with age. More than 65% of bladder cancer cases are subjects older than 65 years [1; 5]. For example, in Canada, the ten-year probability (expressed as a percentage) of developing bladder cancer for a male is 0.1, 0.2, 0.7, 1.3, and 1.5 for age groups (in years) of 40-49, 50-59, 60-69, 70-79, and 80-89, respectively. For women, the corresponding probabilities are 0.04, 0.1, 0.2, 0.3, and 0.4, respectively [6].

2.3.2. Tobacco Smoking

Cigarette smoking is a well-established risk factor for human bladder cancer. According to several review articles, the strong causal link has consistently been demonstrated in many epidemiologic studies (cohort, case-control, and cross-sectional study designs) [1-3; 5; 7; 22-26; 42]. Almost all studies have documented the existence of dose-response relationship between the amounts of cigarette smoking and risk of bladder cancer across sex, age or race groups. The relative risk estimates from the majority of the studies range from about 2.0 to 5.0 (smokers vs. non-smokers). The magnitude of risk depends on the intensity and

duration of smoking [24]. Many studies have also shown that current smokers are at increased risk of developing bladder cancer as compared to former smokers. Some of the studies found that risk of bladder cancer decreased following the cessation of smoking [22; 44]. Smokers of black-tobacco cigarettes are at higher risk than smokers of blond-tobacco cigarettes [26].

In the Western world, cigarette smoking is responsible for the occurrence of 25-30% and 50-60% of all bladder cancer cases in women and men, respectively, rendering it the most important risk factor for bladder cancer [3; 5; 44]. About half of all male bladder cancer deaths in the United States are attributable to cigarette smoking. Amongst women, cigarette smoking is responsible for about 37% of bladder cancer deaths [2].

The exact biologic and carcinogenic mechanism of cigarette smoking in relation to bladder cancer is not yet determined [5; 25]. It is thought that aromatic amines (2-naphthylamine and 4-aminobiphenyl), the agents known to cause bladder cancer, are present in tobacco smoke and are responsible for triggering the carcinogenic process in the human bladder [25; 44]. This mechanism could be supported by the fact that black tobacco has higher concentrations of aromatic amines than blond tobacco and that smokers of black-tobacco cigarettes are at higher risk than smokers of blond-tobacco cigarettes [44]. It has also been hypothesized that there may be some additional factors that could modify smoking-related risks of bladder cancer [25]. For example, some ethnic groups (e.g., Polynesian men and African-Americans) experience lower incidence of bladder cancer than whites despite higher rates of smoking. The factors leading to these

differences are thought to be related to metabolism of smoking-related carcinogens and exogenous factors such as diet or vitamin intake [25; 44].

2.3.3. Occupational and Chemical Exposures

In 1895, Dr. Rehn described 3 cases of bladder cancer in subjects who had worked in the German chemical dye industry. Based on his findings, Dr. Rehn suggested a role for occupational chemicals in the etiology of bladder cancer. Later, large-scale studies conducted amongst industrial workers in the dye, rubber, textile, leather, and aluminum industries, confirmed the findings of Dr. Rehn, and identified 2-naphthylamine, benzidine, 4-amino-biphenyl (4-ABP), aniline, and other aromatic amines as carcinogens responsible for the development of human bladder cancer. For example, one important cohort study showed a 20-fold excess of bladder cancer in textile dye industry workers compared to the general population [50]. Presently, benzidine, 4-ABP, and 2-naphthylamine are banned for use in the work place [5; 23; 44]. In the 1990s, additional aromatic amines and related chemicals (4,4-methylene 2-chloroaniline and o-toluidine) associated with the development of bladder cancer, were identified [3]. Based on the findings of many epidemiologic and occupational studies, many occupations have been declared as high-risk occupations. Especially strong associations are found for workers in the dye, rubber, and chemical industries. Other well-established or suggested high-risk occupations are leather workers, painters, drivers of trucks, buses or cabs, aluminum, steel or iron workers, hairdressers, printers, dry cleaners, carpenters, cooks, tailors, petroleum workers, clerical workers, welders, railroad workers, food processing workers, and butchers. Yet for many of these occupations, specific chemicals responsible for the development of bladder cancer have not been identified [5; 23; 44]. Associations between the

duration of employment and risk of bladder cancer, found in many studies, support a dose-response relationship [5; 44].

It has been estimated that about 20-25% and 11% of all bladder cancers in men and women, respectively, are attributable to occupational factors [44; 51-52].

2.3.4. Dietary Factors (Nutrition)

A role of diet in the development of human bladder cancer is plausible. It is well known that many food substances and their metabolites are excreted through the urinary tract. Some of these metabolites may be carcinogens. Some studies have also shown that supplementation with dietary retinoids inhibited bladder carcinogenesis in experimental animals. Epidemiologic evidence on the effects of dietary factors in relation to bladder cancer risk is not very convincing. The majority of the small number of studies that investigated fruit and vegetable consumption found inverse associations between the levels of fruit and vegetable consumption and risk for bladder cancer. Moreover, intake of carotenoids and Vitamin A has been associated with a decreased risk of bladder cancer. Some studies have demonstrated a positive relationship between fat intake and risk of bladder cancer [5; 7; 44]. One meta-analysis reviewed 38 articles (case-control and cohort studies) reporting the associations between dietary factors and risk of bladder cancer [53]. Elevated risks were found for diets low in fruit intake (RR = 1.40, 95% CI: 1.08, 1.83). Diets, low in vegetable intake, were related to a slightly increased risk (RR = 1.16, 95% CI: 1.01, 1.34). This study has also shown that those with a diet high in fat intake, compared to those with a diet low in fat intake, were at increased risk of bladder cancer (RR = 1.37, 95% CI: 1.16, 1.62).

2.3.5. Coffee, Alcohol Consumption, and Artificial Sweeteners

Many epidemiologic and meta-analytical studies have assessed the relationship between coffee consumption and risk of bladder cancer. In general, the findings are inconsistent. Only few of the positive findings exhibited dose-response relationship [5]. In human studies positive and weak relationships have been found more often.

The accurate measurement of coffee consumption in epidemiologic studies is difficult. The problems related to an accurate measurement may lead to exposure misclassification of the study subjects with respect to levels of coffee consumption. Another concern is that observed positive relationships between coffee intake and bladder cancer risk might partially or entirely be due to inadequate control for smoking status, resulting in residual confounding [5; 7; 44]. This makes the interpretation of study results even more difficult.

Two systematic reviews attempted to pool the results from case-control studies on the association between coffee drinking and the risk of bladder cancer [54-55]. The first study found a weak positive association between coffee drinking and bladder cancer risk only in men (4-5 cups per day vs. non-drinkers: OR = 1.16, 95% CI: 1.05, 1.28) [54]. The second study analyzed only non-smokers in order to avoid the presence of residual confounding by smoking [55]. An increased risk associated with the consumption of ten or more cups per day was observed for both women (OR = 2.0, 95% CI: 0.9, 4.4) and men (OR = 1.8, 95% CI: 0.7, 4.3). However, none of the estimates reached conventional levels of statistical significance. The authors of this review concluded that heavy coffee drinkers had a small excess risk relative to non-drinkers which could not be attributed to

confounding by smoking. In summary, based on the evidence overall, which is not so strong, one could suggest that if coffee is a carcinogenic substance to human bladder, then it is a weak one.

Alcohol consumption as a risk factor for bladder cancer has also been investigated. The results are not consistent. However, the majority of the evidence has indicated that there was no association between alcohol intake and bladder cancer risk. One meta-analysis of observational studies showed a slightly elevated, although statistically non-significant risk of male bladder cancer. The age- and smoking-adjusted summary odds ratio for current alcohol drinkers compared to non-drinkers was 1.3 (95% CI: 0.90, 2.00) [56]. In a hospital-based case-control study conducted by Pelucchi et al., an estimate of OR for alcohol drinkers ($\geq$6 drinks/day), relative to non-drinkers, was 0.84 (95% CI: 0.58, 1.22). The risk of bladder cancer in wine drinkers ($\geq$ 5 drinks/day) was not statistically significantly different from that in non-drinkers (OR = 0.86, 95% CI: 0.60, 1.23) [57].

It has been suggested that the positive findings regarding the effects of alcohol consumption on bladder cancer risk could reflect chance or residual confounding [5; 44; 56].

Animal experiments have demonstrated that male rats fed with high doses of saccharine or cyclamate in utero developed bladder cancer. This finding could not be confirmed by animal studies using postnatal exposure to saccharine in rats [5]. Human studies looking at maternal intake of saccharine during pregnancy and the risk of bladder cancer in the offspring have not yet been conducted. The evidence based on epidemiologic studies has

indicated that the consumption of artificial sweeteners is not associated with risk of bladder cancer. For example, one large population-based case-control study (3,000 cases and 5,700 controls) observed a relative risk of 1.02 (95% CI: 0.92, 1.11) for ever users versus never users of artificial sweeteners [7]. Based on the reviewed evidence, results from most case-control studies exploring the association between the consumption of artificial sweeteners and risk of bladder cancer, have been negative [5; 7; 24; 44].

2.3.6. Chlorination By-Products and Arsenic in Drinking Water

Some studies have suggested that the presence of chlorination by-products or arsenic in drinking water might be associated with an increased risk of bladder cancer [5; 7; 44]. One meta-analysis systematically reviewed six case-control and two cohort studies of chlorination by-products and bladder cancer [58]. The summary OR estimates indicated that men who had ever consumed chlorinated drinking water had a slightly increased risk of bladder cancer (OR = 1.4, 95% CI: 1.1, 1.9), but women did not have any increased risk (OR = 1.2, 95% CI: 0.7, 1.8). Two case-control studies investigated the association between the duration of chlorinated water use and risk of bladder cancer [59-60]. The first study demonstrated that those who used chlorinated surface water for 35 years or longer had 1.41 times the risk of those whose period of exposure was under 10 years (OR = 1.41, 95% CI: 1.09, 1.81) [59]. The second study showed a dose-response relationship between the duration of exposure to chlorinated water (surface and ground) and risk of bladder cancer in men. Specifically, the ORs for chlorinated surface water use were 1.0 (referent), 1.1, 1.3, 1.5, and 1.9 for 0, 1-19, 20-39, 40-59, ≥ 60 years of exposure, respectively [60].

Between 1989 and 1999, the U.S. Environmental Protection Agency (EPA) conducted a risk assessment for arsenic in drinking water and found sufficient evidence to conclude that ingestion of arsenic in drinking water causes skin, bladder, and lung cancer [7]. Before EPA initiated its investigation, two studies had examined the association between the concentration of arsenic in drinking water and bladder cancer-specific mortality rate. Both studies found positive associations [61-62]. Specifically, Wu et al., used direct standardization to calculate age-adjusted mortality rates from bladder cancer for groups of the Taiwanese population in an endemic area of blackfoot disease, who ingested water with different levels of arsenic. The calculation of mortality rates by sex were based on the 1976 world population. For both men and women, significant dose-response relationships were observed between arsenic levels in well water (Natelson's method for the measurement of water arsenic concentration) and mortality rates from bladder cancer. In men, the age-adjusted mortality rates were 22.64, 61.02, and 92.71 (P value for trend < 0.01) for < 0.30, 0.30 - 0.59, and > 0.59 ppm of water arsenic concentrations, respectively. In women, the corresponding rates were 25.60, 57.02, and 111.30 (P value for trend < 0.01) [61]. The other study used indirectly standardized mortality ratios (SMR) in order to compare bladder cancer mortality rate in an endemic area of blackfoot disease to that in the general population of Taiwan (the 1971 population of Taiwan as standard). Those who used water from artesian wells with higher concentrations of arsenic experienced a higher bladder cancer mortality rate compared to the general population of Taiwan (Men: SMR = 11, 95% CI: 9.33, 12.67; Women: SMR = 20, 95% CI: 17.02, 23.16) [62].

2.3.7. Fluid Intake

Bladder urothelium is in direct contact with carcinogens excreted in the urine. Based on this, it has been suggested that increased consumption of fluids may reduce the exposure to these carcinogens by diluting the urine and shortening the period of contact due to increased frequency of urination [5]. According to the reviewed evidence, the results stemming from case-control and cohort studies regarding the protective effect of high fluid intake are conflicting [5; 7; 63-64]. For example, one American cohort study found that men with the highest fluid intake had half the risk of bladder cancer compared to those with the lowest intake (RR = 0.51, 95% CI: 0.32, 0.80). The protective effect was observed for water as well as for other beverages [63]. In contrast, a Dutch cohort study obtained negative results suggesting that total fluid intake was not associated with the risk of bladder cancer in men and women (RR = 0.91, 95% CI: 0.65, 1.29) [64]. One large population-based case-control study, conducted in the United States (Utah), found no relationship between total fluid intake and bladder cancer risk [65].

Difficulties in measuring total daily fluid intake accurately and reliably could contribute to obtaining these conflicting results [7]. Another explanation for the inconsistent results across studies is that in some regions drinking water contains relatively high levels of chlorination disinfection by-products effect of which could offset the benefit of high volume fluid intake (reduced risk for bladder cancer) or simply increase the risk. Because drinking water with high levels of chlorination by-products has an unpleasant taste, certain proportion of population will consume bottled water. Given this, two studies

conducted in the same region but failing to account for this factor could yield different results [63].

2.3.8. Urinary Tract Diseases

Schistosoma haematobium has been linked to bladder cancer, but only to squamous cell carcinoma, not to transitional cell carcinoma – the predominate type of bladder cancer in North America. Especially high rates of squamous cell carcinoma of the bladder are observed in the Middle East and some parts of Africa, where schistosomiasis is a wide-spread problem (endemic areas)[7]. The link between *schistosoma haematobium* and squamous cell carcinoma of bladder is not well understood. It is thought that eggs of the parasite are deposited in the bladder wall and urine, causing chronic irritation of urothelium, cystitis, and haematuria. The chronic inflammatory process is hypothesized to result in increased cell proliferation and neoplastic changes in the urothelium, leading to squamous metaplasia [24; 44]. Other hypotheses have suggested that *schistosoma haematobium* parasites could produce carcinogenic nitroso compounds in the patient's urine, or the parasite's antigen might depress the immune system of infected patients [5; 44].

Certain medical conditions (i.e. prostatic hypertrophy, kidney stones) leading to urinary stasis may also increase the risk of developing bladder cancer. One review article reported a Swedish cohort study in which 61,114 patients hospitalized for kidney stones were followed up by means of record linkage for 18 years. The age- and sex-standardized incidence ratio was statistically significant, indicating an elevated risk of bladder cancer for the patients diagnosed with kidney stones, compared to the general population (SIR = 1.4, 95% CI: 1.3, 1.6)[7].

2.3.9. Drugs and Treatment-Related Carcinogenesis

A number of pharmaceutical agents have been linked to bladder cancer. The strongest evidence concerns high consumption of phenacetin-containing analgesics. Based on the results of several case-control studies, the estimates of relative risk for ever users of phenacetin versus non-users ranged from 1.5 to 6.5 [5; 7].

Case reports and case-series have suggested that cyclophosphamide, an alkylating agent used for the treatment of tumours, may be associated with an increased risk of bladder cancer [5]. In a Danish prospective study, patients with non-Hodgkin's lymphoma treated with cyclophosphamide had about a sevenfold risk of bladder cancer compared to those who had not been treated with cyclophosphamide (RR = 6.8, 95% CI: 3.2, 14.2) [66].

Exposure to radiation may also increase the risk of bladder cancer. A few studies, summarized in one review article, have shown that female patients who underwent therapeutic pelvic radiation for uterine bleeding were at two-to-four fold increased risk of bladder cancer [5].

2.3.10. Family History

Based on the reviewed evidence, results from a few case-control studies suggest the existence of familial predisposition to bladder cancer. In one case-control study, an odds ratio of 2.50 (95% CI: 1.10, 5.80) was found for positive family history in males compared to matched controls [44]. One large case-control study even found that the

familial risk of bladder cancer was especially high among heavy smokers, suggesting the presence of gene-environmental interactions [67].

2.3.11. Miscellaneous Factors

The follow-up study of 75,991 atomic bomb survivors in Japan (between 1950 and 1985) found an increased bladder cancer-specific mortality rate amongst those exposed to a higher dose of ionizing radiation (> 0.005 Gy), as compared to those exposed to a lower dose of ionizing radiation (≤ 0.005 Gy) (RR = 2.13, 90% CI: 1.40, 3.28) [68].

2.4. Measurement/Assessment of Physical Activity in Epidemiologic Research

Physical activity may be defined as any bodily movement produced by skeletal muscles that results in energy expenditure [29]. Physical activity, depending on its purpose, may be divided into three main groups: occupational (work-related), household (garden and other household activities), and leisure-time or recreational (exercise and sport) [28]. Depending on the study question, researchers use leisure-time (or recreational) activity separately or combined with occupational physical activity [69-70].

The methods for measuring physical activity/energy expenditure include: laboratory (direct/indirect calorimetry), physiological (doubly labelled water, resting heart rate, and maximal oxygen uptake VO_2), and field methods (questionnaires). Although laboratory and physiological methods measure physical activity/energy expenditure more accurately than questionnaires, they are not widely used in epidemiologic studies due to high costs and restricted applicability [69-70].

Epidemiologic research for the assessment of physical activity often employs field methods, chiefly self- or interviewer-administered questionnaires. The responses on the questionnaires are used to estimate the physical activity level/energy expenditure for the individual. Specifically, the responses are counted and converted into (kilo-) joules per day or week, (kilo-) calories, in hours (duration), metabolic equivalents (MET-values), or as a summary activity score [69].

The questionnaires may elicit general patterns or precise details of physical activity over a given period of time. There are three main groups of physical activity questionnaires: global, single-item, and comprehensive. Global questionnaires differ from single-item or comprehensive questionnaires in a way that they assess an individual's level of physical activity in relative terms, whereas single-item and comprehensive questionnaires summarize actual levels of physical activity in absolute terms. Global and single-item questionnaires are easy to use and they allow relatively quick assessment of general patterns of physical activity. For example, a global questionnaire may ask a study participant to compare his/her activity level to that of other people of the same sex or age. The responses could fall in the following categories: much less, less, the same as, more, or much more than others (men or women of the same age). A single-item questionnaire can measure weekly (or monthly) frequency of moderate and vigorous physical activity. A comprehensive questionnaire may assess the time spent in sleep, moderate, hard, and very hard physical activities. This type of questionnaire can simultaneously elicit the physical activity information on work and leisure-time activities. The frequency of physical activity is measured by recording the number of times per week, month or year,

or the number of months per year that a particular activity was performed. In some cases, the study subjects are simply grouped with respect to session frequency categories (often, very often, and always) [28].

Ideally, a questionnaire should address four dimensions of physical activity: type (occupational, recreational, or household), frequency (e.g., number of days per week), duration (e.g., number of hours per day), and intensity (light, moderate, and strenuous/vigorous). Epidemiologic studies use a wide range of methods to measure the dimensions of physical activity. Depending on the availability of data, researchers categorize study subjects with respect to levels of physical activity defined by different combinations of duration, frequency, type, and intensity. If data are available on all dimensions of physical activity (type, frequency, duration, and intensity), then a summary score is calculated by combining data on the frequency, duration, and intensity of each specific type of activity [28]. The information on type, duration, frequency, and intensity of activity, elicited by a questionnaire, is usually converted into units of energy expenditure by using tables published in literature [69]. Often the session frequency and/or duration (in minutes) are weighted by the intensity scores for each type of activity in order to calculate a summary physical activity score for all the performed activities (for example, expressed as the number of MET-hours per year/month/week) [28].

The intensity of physical activity is a very important component, indicating the amount of energy expenditure or oxygen consumption per unit of time. Specific physical activities differ in terms of their intensity (e.g. walking vs. running). Often researchers employ

different coding systems for expressing the intensity of physical activity. These differences, in turn, limit the comparability of results across studies. One common way to express the intensity of a specific activity is to assign a metabolic equivalency threshold (MET) value. One MET is an amount of energy expended while sitting in a rested state, and is equal to 3.5 ml of oxygen consumed per kilogram body weight per minute or 1 kilocalorie per kilogram body weight per hour. After assigning a MET value to each specific activity, the intensity is then expressed in relative terms, as the ratio of the associated metabolic rate for the specific activity divided by the resting metabolic rate [28; 69].

Some researchers assess work-related physical activity by using occupational status (based on job title) or work-related activity as a surrogate measure to estimate levels of physical activity for a certain point or time period [70]. A common method is to compare separate risks of a health-related outcome (e.g., cancer) for two or more different occupational groups (by specifying current and/or longest-held job over the person's lifetime) relative to a reference category requiring the least amount of physical activity (e.g., housewife). Other studies have investigated the relationship between occupational physical activity and cancer risk in a single occupational setting (industry), by comparing risks of office workers (sedentary activity group) to those of workers with other job-related duties requiring light, medium, or heavy activities. Some researchers have grouped the study subjects according to the duration of physical activity (the percentage of the time as sedentary, moderate, and/or rigorous/strenuous), required by any specific job title [28].

The number of years and types of physical activity as well as total lifetime physical activity may be as important in the etiologic process of any given disease as changes in physical activity over time. If the elicited information does not pertain to an etiologically relevant time period, a questionnaire even though reliable and/or valid, may fail to detect an existing important effect of physical activity. Therefore, prospectively recorded and gathered data on all four dimensions (type, frequency, duration, and intensity) throughout the subject's lifetime would provide crucial information in assessing the cumulative influence of physical activity on risk of a disease. The best approach would be to measure physical activity repeatedly over many years with a follow-up for incident cases of a disease. But this approach is often costly and not feasible [70].

As an exposure, physical activity is a complex and multidimensional variable, which in its entirety has a wide range and is not constant over one's lifetime. Therefore, the accurate measurement of physical activity in epidemiologic investigation is challenging and problematic. The presence of these methodological difficulties limits the interpretability of some studies [27-28; 69-70].

The methods used to estimate physical activity are subject to systematic, random, or both (simultaneous occurrence) types of errors. For example, self-report of physical activity may be inaccurate because of difficulties in recall of the performed activities in previous years. Besides, the perception of intensity and duration of physical activity may be subjective [27]. In general, whether the recall is for a recent or earlier period, it is more accurate for strenuous physical activity than for moderate-intensity activity [28; 69].

It is very important to validate/test a questionnaire by estimating the magnitude of error, i.e., assess its degree of reproducibility and validity in a sample of subjects who would be representative of the final study sample.

In epidemiologic research the most common method to determine the degree of reproducibility is based on test-retest of the same questionnaire over a pre-specified period of time. The estimated reproducibility of any physical activity questionnaire varies with the time period (if the period is too short, the subject's recall may influence his response; if the period is too long, changes may occur in activity patterns) between two or more administrations of the same questionnaire and the wording of questions.

If a questionnaire systematically under- or over-estimates the true levels of physical activity, its validity should be questioned. The assessment of the questionnaire's validity (criterion/concurrent validity) involves finding an appropriate 'gold standard' method. However, it is hard to find a 'gold standard' for physical activity measurement. This is related to high costs, different constructs (physical activity versus physical fitness) and units of measurement [69]. The physical activity questionnaires used in recent epidemiologic studies have been shown to correlate with oxygen uptake, resting pulse, treadmill fitness tests, and the results of other physical activity surveys. The majority of correlation coefficients were 0.50 or greater [28].

2.5. Measurement of Body Weight in Epidemiologic Research

Body fat is an important health determinant. The measurement of body fat in epidemiologic studies is very difficult. The laboratory standard involves 100% immersion in water and is clearly not feasible in population-based studies [28].

Body weight has been used in epidemiologic studies as an indirect measure of body fat mass. However, body weight is not a good measure of body fat because it is confounded by height and muscle bulk. A number of measures have been proposed to estimate body fat mass (change in weight, waist-to-hip ratio, skinfold and waist circumference). However, the most commonly employed measure is the Quetelet or body mass index, calculated by dividing the weight in kilograms by the square of the height in meters (kg/m^2). Adjusting for height, by calculating body mass index (BMI), removes variation in weight due to height, and renders BMI as a better indicator of body fat mass. But because BMI cannot differentiate muscle mass from fat mass, it is not an ideal measure of body fatness or adiposity. Body mass index has been validated as a measure of body fat mass, showing a strong correlation ranging from 0.82 to 0.91 [28].

The body mass index should be based on measured height and weight. However, epidemiologic studies often use self-reported weights. Despite the tendency to over- or under-report weight, self-reported weights have been shown to correlate highly with measured weights (usually, $r > 0.95$), and to be consistently associated with diseases that are known to be related to excess adiposity [28]. Self-reported weights, referring to periods many years in the past, retain a high degree of validity [28; 71].

Although it is easy to measure study subjects' body weight or BMI, investigators should be aware of the potential for reverse causality. Retrospective studies may be prone to this kind of bias. If anthropometric measures are collected around the time of diagnosis and treatment, it would be hard to determine whether these measurements were consequence or cause of the disease of interest. For example, in a case-control study, if cases reported lower weight or BMI, compared to controls, at the time of interview, it would not be clear whether low body weight/BMI led to increased risk of the disease or low body weight/BMI arose as consequence of the disease (e.g. cachexia due to cancer).

2.6. Physical Activity, Body Weight and Cancer Risk: Possible Roles of Endogenous Hormones and Immune System

Substantial epidemiologic evidence, accumulated over the past two decades, suggests that an excess body weight and physical inactivity are associated with increased risk of some types of cancers. For example, the International Agency for Research on cancer (IARC)-working group estimated that about 11% of all colon cancers are attributable to high BMI and 13-14% to physical inactivity [28].

Based on the results obtained from many studies, it appears that effects of physical activity and body weight on cancer risk are independent of each other [27; 73]. However, since these two factors are closely related, it is also possible that the protective effect of physical activity against cancer operates through inducing a body weight reduction: regular or high levels of physical activity might lead to the reduction of body fat mass.

Depending on the study hypothesis, adjustment of the effect of physical activity for indices of body weight (or fat) may not be appropriate [27].

One hypothesis is that the effects of physical activity and body weight on cancer risk might operate through alterations in the metabolism of endogenous hormones (sex steroids, insulin, and insulin-like growth factor-I) [27-31; 72-73]. Another mechanism suggests that the protective effect of physical activity against cancer might be mediated by changes in body immune system capacity [27].

2.6.1. Physical Activity and Cancer Risk

The evidence based on the findings of many case-control and cohort studies of men and women suggests that both leisure-time and occupational physical activity protect against overall cancer risk [29]. Epidemiologic findings regarding the protective effects of regular physical activity are relatively consistent for colon, postmenopausal breast, endometrial, prostate, and esophageal (adenocarcinoma) cancers [27-29]. For example, an average risk reduction for colon cancer ranged from 40 to 50% when the highest and lowest levels of physical activity were compared. Most studies of colon and breast cancer have found that these protective effects could be achieved either through increased intensity, duration or frequency of physical activity. The majority of the studies focusing on breast cancer demonstrated 20 – 40% risk reduction in the most physically active women relative to those with non-active or sedentary life-styles [28-29].

2.6.2. Body Weight (Body Mass Index) and Cancer Risk

Based on the detailed review conducted by the IARC group, there was sufficient evidence to suggest that excess body weight was associated with an increased risk of cancers of the colon, breast in postmenopausal women, endometrium, kidney (renal cell tumours), and adenocarcinomas of the esophagus [28].

Amongst women with a BMI above 25 kg/m^2, an excess body weight has been found to be associated with two- or three-fold increased risk for endometrial cancer. In most epidemiologic studies (case-control and cohort) of body weight and colon cancer (men and women), the relative risk for those in the highest versus the lowest quartiles/tertiles of BMI, has been shown to range from 1.4 to 3.0 [27-29]. The majority of the colon cancer studies found that those with BMI > 30 kg/m^2 had about a double risk of those with BMI < 23 kg/m^2 [28].

Besides demonstrating the harmful effect of an excess body weight, measured at one point in time, adult weight gain over time has also been found to be an important predictor for breast and colon cancers [28; 73].

The results obtained from studies investigating body weight and cancer occurrence are more convincing and consistent than those on physical activity. This could partially be due to the well-known difficulties involved in measuring a multidimensional variable such as physical activity [72].

2.6.3. The Role of Endogenous Hormones

The most important biologic mechanism hypothesized to link physical activity, body weight, and risk of cancer, involves the modulation of endogenous hormones. Although the exact pathway between physical activity, body weight and cancer risk is not known, it has been postulated and shown that physical inactivity and excess body weight may lead to insulin resistance and increased secretion/circulation of such endogenous hormones as plasma insulin, sex hormones (estrogens and androgens), and insulin-like growth factor (IGF-I) [27-31]. The excess of insulin in blood is related to the decreased levels of insulin-like growth factor-binding proteins (IGFBPs) and sex-hormone binding globulin (SHBG). The decreased levels of insulin-like growth factor-binding proteins and sex-hormone binding globulin lead to the increased levels of insulin-like growth factor and free sex hormones. The chronic exposure to high levels of insulin-like growth factor and sex hormones is implicated in enhancing carcinogenesis [27-31; 72-73].

2.6.4. The Role of Immune System Capacity

Another proposed mechanism suggests that the effects of physical activity on risk of cancer may operate through immune function. The immune system has a capacity of detecting and eliminating neoplastic cells [27]. It has been postulated that physical activity increases the numbers and activities of circulating natural killer cells such as lymphocytes, granulocytes, monocytes, and macrophages, which in turn may inhibit the carcinogenic process [32]. The magnitude and direction of physical activity-related immune response in relation to cancer is not clear but it is thought to depend on exercise type, intensity, and duration [27]. The role of natural killer cells in tumour immunity has been

demonstrated in animal studies indicating that exercise-induced enhancement of innate immunity may have a protective role [32]. Further research is warranted to elucidate the potential role of physical activity-induced enhancement of immune function in relation to the reduction of cancer risk.

2.7. Endogenous Hormones in Relation to Bladder Cancer

Results of many studies indicate the involvement of endogenous hormonal factors in the carcinogenesis of bladder cancer [8-22]. For example, unexplained male excess risk for human bladder cancer might be due to the sex-based differences in the exposure to endogenous sex hormones [8-10]. Other retrospective studies on humans have shown that sex hormones or their proxies (reproductive factors) could modify bladder cancer risk [10; 12; 19; 21]. For example, Cantor et al., observed a decreased risk of bladder cancer in women with one or more live births relative to nulliparous women (OR = 0.51, 95% CI: 0.30, 0.88) [10]. One study found that women with bladder cancer had lower levels of progesterone and a shorter reproductive period than controls [19]. Another case-control study in Italy found that women who had ever used hormone-replacement therapy had an increased risk of bladder cancer relative to those who had never used it (OR = 3.29, 95% CI: 1.49, 7.25) [22]. A study in Finland looked at the expression of androgen receptor protein as a possible marker of androgen action in the cystectomy specimens of human bladder cancer. About 70% (6 out of 9) of the cases had expressed androgen receptor. Based on the high frequency of androgen receptor expression, the authors of this study suggested a possible role for androgens in the bladder carcinogenesis [12]. Zhao et al., in a nested case-control study, showed that higher levels of insulin-like growth factor were

related to an increased risk of bladder cancer (highest exposure quartile vs. lowest OR = 3.10, 95% CI: 1.43, 6.70) [33]. Three studies investigated the effects of reproductive factors on cancer mortality [74-76]. One of them demonstrated a statistically significant finding indicating that women with at least one live birth, relative to those without any live birth, experienced a lower rate of mortality from bladder cancer: SMR-s across the groups defined by the number of live births 0, 1, 2, 3, and 4 were 2.25, 1.17, 0.59, 0.86, and 0.47, respectively (Test for trend, Chi-square = 4.8) [74]. One should be cautious when interpreting results from mortality studies, since mortality reflects not only the incidence (or risk) of bladder cancer but also survival.

Apart from the studies investigating hormonal factors and the incidence of human bladder cancer, a number of animal studies have shown an effect of endogenous hormones on bladder cancer incidence [11; 14-17; 20]. For example, two studies have demonstrated that rats treated with endogenous sex hormones such as testosterone or estradiol experienced higher bladder cancer incidence than control series of rats [16; 20]. Moreover, castrated male rats experienced a lower incidence of bladder tumours compared to control male rats (approximating the incidence found in female rats). Female rats treated with testosterone had higher tumour incidence than those in control groups. In their study, Reid et al.[14] observed accelerated growth of a transitional cell tumour of the bladder when treated with androgens. Although estrogens suppressed the growth of transitional cell tumour, they appeared to promote the formation of neoplastic squamous cells with a propensity for greater invasion [14].

In summary, the accumulated evidence of animal and human studies suggests that endogenous hormones may be involved in the carcinogenesis of urinary bladder. The majority of studies have shown that male sex hormones androgens (testosterone, dehydroepiandrosterone) enhance the carcinogenesis of bladder cancer [12; 14-15; 20-21]. Zhuang et al. and Imaoka et al., based on their findings hypothesized that androgens may in part explain the excess risk in men compared to women [12; 15]. There is very little evidence regarding the effects of endogenous estrogens [14; 16] and female reproductive factors (as proxies of estrogens) [10; 74-76] on bladder cancer risk. The results of these studies are not consistent. Out of the four human studies [10; 74-76], two demonstrated a harmful effect of nulliparity used as a proxy of increased estrogen exposure during reproductive period [10; 74]. In two animal studies [14; 16], one found that the administration of estrogens (estradiol) was related to the increased incidence of bladder cancer [16]. In the other study the estrogens were shown to promote the growth of squamous cell carcinomas but not transitional cell tumours of bladder [14].

Based on the reviewed evidence, it seems that androgens are more potent carcinogens of urinary bladder than estrogens, and this difference may be in part responsible for higher incidence of bladder cancer in men than in women. Moreover, exposure to different levels of androgens may be an explanatory factor of within sex differences in bladder cancer risk (in men) [12; 14; 15; 21]. The carcinogenic role of estrogens in the development of female bladder cancer should not be excluded (i.e. women with reduced exposure to estrogen may have lower risk of bladder cancer) [10; 16; 22; 74].

2.8. Physical Activity, Body Weight, and Risk of Bladder Cancer

An extensive literature search identified and retrieved 11 epidemiologic studies that investigated the associations of physical activity and/or body weight in relation to risk of bladder cancer [22; 35-41; 77-79].

Six studies (two case-control and four cohort) examined the link between physical activity and risk of bladder cancer [35-38; 40; 79]. The two case-control studies investigated the relationship between occupational physical activity and risk of various cancers in males [40; 79]. Both studies used hospital-based controls diagnosed with cancer or other non-cancerous conditions. Subjects in both studies were categorized with respect to their physical activity levels based on occupational job codes. The results from both studies, based on the age- and smoking-adjusted ORs, indicated that levels of occupational physical activity were not associated with risk of bladder cancer in men. Specifically, men with low (or moderate) occupational physical activity levels had about the same risk of bladder cancer as those with high levels of occupational activity (OR = 1.0; 95% CI: 0.6, 1.5 and OR = 1.1; 95% CI: 0.8, 1.5). The number of bladder cancer cases analyzed in these studies ranged from 267 to 1,080.

Of the four cohort studies [35-38], three found no association between levels of recreational physical activity and bladder cancer risk [35-36; 38]. In two studies [35-36], men in physically active groups had similar risks for bladder cancer to those in non-active groups (age-adjusted RR = 0.72, P-value = 0.30 and age-, BMI-, smoking-adjusted RR = 1.05, 95% CI: 0.60, 1.84). The number of cases in these studies ranged from 58 to 70 [35-36]. Similarly,

in a follow-up study of postmenopausal women, Tripathi et al., found no evidence for an association between leisure-time physical activity and risk of female bladder cancer. This study was based on 112 newly diagnosed bladder cancer cases. The age-adjusted RR for bladder cancer amongst women in 'high' activity group relative to 'low' activity group was 0.73 (95% CI: 0.46, 1.17) [38]. In contrast, one cohort study found that the level of physical activity was associated with risk of male bladder cancer [37]. Wannamethee et al., based on 92 incident cases of bladder cancer, observed a statistically significantly increased risk of bladder cancer in men who engaged in vigorous physical activity as compared to those who were physically inactive (age-, smoking-, BMI-, alcohol intake-adjusted RR = 2.06, 95% CI: 1.08, 3.97). The authors also demonstrated a marginally statistically significant dose-response relationship between the levels of physical activity and bladder cancer risk (test for trend=0.04) [37].

Two Scandinavian cohort studies [41; 78] employed a method of indirect standardization to compare the cancer experience in cohorts (men and women) of subjects discharged from hospitals with a diagnosis of obesity to that in the general population. In both studies age-standardized incidence ratios (SIRs) and 95% CIs were computed. The results are conflicting [41; 78]. Moller et al., who ascertained 64 male and 35 female bladder cancer cases, found an obesity-related excess risk for bladder cancer in men (SIR = 1.3, 95% CI: 1.0, 1.6), but not in women (SIR = 1.1, 95% CI: 0.7, 1.5) [78]. In contrast, Wolk et al., based on 32 male and 35 female bladder cancer cases, demonstrated that the study cohort of male subjects (men with discharge diagnosis of obesity; BMI > 30.0 kg/m^2) experienced similar bladder cancer incidence as the general population of men (SIR =

1.1, 95% CI: 0.8, 1.6). As for women, the study cohort (women with discharge diagnosis of obesity; BMI > 28.6 kg/m^2), experienced an elevated bladder cancer incidence relative to the general population of women (SIR = 1.4, 95% CI: 1.0, 1.9) [41].

Three case-control studies have also been done [22; 39; 77]. Whittemore et al. investigated the association between weight and bladder cancer risk in former college men. Each of the 106 bladder cancer cases was matched to four randomly chosen classmates (controls) born in the same year and known to be free of bladder cancer at the time of the case's diagnosis. The authors found that a gain of 10 pounds in body weight corresponded, on average, to a 10% increase in bladder cancer risk (age-, college, and class year-adjusted RR = 1.1, 95% CI: 1.0, 1.3). However, the study results did not support the association between BMI and bladder cancer risk [39]. Two hospital-based exploratory case-control studies of bladder cancer examined the association between BMI and risk of bladder cancer in women [22; 77]. Based on the analysis of 110 bladder cancer cases and 298 controls, Pelucchi et al., did not find an association between BMI and bladder cancer risk. The adjusted (for age, study center, and cigarette smoking) estimate of OR for women in the highest BMI tertile (BMI > 27.25) versus the lowest tertile (BMI < 23.75) was 0.61 (95% CI: 0.33, 1.14) [22]. Kabat et al. [77] in a matched case-control study investigated various characteristics in relation to bladder cancer risk amongst non-smoking men and women. Seventy-six male and 76 female cases of bladder cancer were matched to 238 male and 254 female controls, respectively. The matching variables were sex, age (within 5 years), race, hospital, and year of interview. The authors did not find any association

between body mass index and bladder cancer risk (ORs and 95% CIs were not reported) [77].

In summary, for both men and women, the results from studies reporting relationships between body weight and bladder cancer risk, are not consistent [22; 39; 41; 77-78]. As for the studies reporting the associations between levels of physical activity and risk of bladder cancer in men, four out of five found no association [35-36; 40; 79]. There was only one study that investigated the association between physical activity and bladder cancer in women, results of which indicated that physical activity was not associated with risk of bladder cancer [38].

An epidemiologic study exploring simultaneously physical activity, relative body weight, and risk of bladder cancer, as main exposure and outcome variables, has not been previously reported [42].

2.9. Methodological Limitations of Previous Studies

The approaches used by the above-mentioned studies for the specific investigation of the association between physical activity, body weight, and risk of bladder cancer were methodologically deficient or inadequately rigorous. Many of the studies had limitations in the adequate measurement of physical activity. A common problem was that the physical activity indices were too crude to appropriately classify the subjects with respect to their daily activities. For example, two studies did not have direct measurements of physical activity levels, but instead they used occupational codes as proxies to the

subjects' activity levels. Both studies used relatively crude categories (low or sedentary, moderate, and high or active) of physical activity levels [40; 79]. Two cohort studies used physical activity indices that did not include one important component - duration of each performed activity [37-38]. One cohort study simply dichotomized the study subjects based on whether or not they had been engaged in sports as college students (cut-point 5 hours per week) [35].

Another methodological limitation common to these studies was inadequate control for confounding of the relationship of interest. For example, the studies which did not focus exclusively on bladder cancer as the outcome were not able to adjust the associations of interest for occupation - an important risk factor of bladder cancer [35-37; 40-41; 78-79]. Since the two Scandinavian cohort studies calculated SIRs, their ability to control for confounding was limited to age and sex only [41; 78]. Only three studies controlled for smoking adequately by considering smoking duration, number of cigarettes smoked per day (in packs), and smoking status (never, former or current) [22; 38-39]. One study, in order to adjust for smoking, used the number of cigarettes smoked per day only for current smokers but not for former smokers [37]. Another study adjusted for smoking through crude dichotomization of smoking status (never vs. ever) [40]. In only three of the 11 studies were physical activity and body weight adjusted for each other [36-38].

Selection bias may have been present in the case-control studies that used patients with certain types of cancer (or other conditions) as controls (hospital-based) [40; 79]. The direction of the resultant bias could be towards the dilution of the effect estimates.

Sample-size limitation is also a problem. The majority of studies based their analyses and conclusions on very small numbers of bladder cancer cases. For example, the sample size of six studies ranged from 32 to 106 male cases [35-37; 39; 41; 78]. The conclusions on women were also based on small numbers of bladder cancer cases, ranging from 35 to 112 [22; 38; 41; 78]. Small numbers are known to result in unstable effect estimates and low study power [42].

3. Research Objectives

Physical activity and/or body weight may exert their effects on bladder cancer risk through hormonal (serum insulin, androgens, estrogens, and insulin-like growth factors), immune or yet other unknown causal pathways. Based on the well-known differences in hormonal milieu between men and women, it is plausible that the effects of physical activity and body weight on bladder cancer risk are mediated differently in men and women. Moreover, given the findings of several studies between female hormonal factors (endogenous or exogenous) and bladder cancer risk, one may hypothesize that female reproductive factors, taken as surrogate measures of exposure to endogenous estrogens, are independently (of physical activity and body fat mass) related to bladder cancer risk in women.

3.1. The Study Hypotheses

The underlying hypothesis of this study is that physical activity and/or body weight are related to risk of bladder cancer. The effects of physical activity and/or relative body weight differ in men and women. There is a relationship between female reproductive factor(s) and risk of bladder cancer.

3.2. The Study Objectives

1. To explore the effects of recreational physical activity and relative body weight (body mass index) in relation to risk of bladder cancer in men and women
2. To determine whether sex is an effect modifier of the effects of recreational physical activity and relative body weight (body mass index) on risk of bladder cancer

3. To examine relationships between female reproductive factors (taken as surrogate measures of endogenous estrogens) and risk of bladder cancer

4. The National Enhanced Cancer Surveillance System: Design

4.1 Overview

This thesis is based on case-control data obtained from the National Enhanced Cancer Surveillance System (NECSS) database. The NECSS is a collaborative effort between Health Canada and the Provincial Cancer Registries. This initiative was funded through the Action Plan on Health and the Environment, Health Canada's part of the Federal Government's Green Plan. The NECSS was developed between 1993 and 1998 to systematically investigate cancer-environment and cancer-lifestyle risk issues in Canada [80].

One of the major components of the NECSS was a case-control surveillance system whose main purpose was the systematic collection of risk factor information from a national population-based sample of newly diagnosed cancer cases and population controls. The objective of this surveillance was to assess the impact of environmental factors on cancer occurrence. The case-control methodology helped to minimize or avoid many limitations that arise in ecological studies. For example, in ecologic studies there are many unmeasured confounding factors that are responsible for the observed patterns on a group level. The ecological fallacy makes it difficult to extrapolate the inferences from the group level to the individual level [80].

The case-control component of the NECSS consists of 20,755 subjects diagnosed with cancer (18 primary sites) in Canada between 1994 and 1997, and 5,039 population controls, randomly selected within the same period. The cases and controls were selected

from the following eight Canadian Provinces: British Columbia, Alberta, Saskatchewan, Manitoba, Nova Scotia, Newfoundland, Ontario, and Prince Edward Island. Information on risk factors from cases and controls was collected using the same protocol. Current military personnel (and their families) and indigenous peoples were excluded from the selection process [80].

4.2. Questionnaire Used for the Collection of Risk Factor Information

A mailed questionnaire (see Appendix A), developed by the former Laboratory Centre for Disease Control (LCDC) and the Provincial Cancer Registries, was used to collect detailed risk factor information from the cases and controls. This information included socio-demographic characteristics (sex, race/ethnicity, age at interview, education, income, and marital status), life-time residential history (up to 1997), occupational history, anthropometric measures (height and weight), lifestyle (recreational physical activity, smoking history, and alcohol consumption), dietary habits (meat, vegetable, vitamin/mineral supplements, and coffee consumption), and reproductive history (age at menarche, menopausal status, age at menopause, age at first pregnancy, and parity) [80].

4.3. Case Recruitment

Between 1994 and 1997, the Provincial Cancer Registries identified subjects diagnosed with 18 primary cite-specific cancers (prostate, breast, colon, leukemia, bladder, kidney, rectum, non-Hodgkin's lymphoma, liver, testis, pancreas, lung, brain, stomach, bone/cartilage, salivary, multiple myeloma, and mesothelioma) through medical/pathology records, and obtained physician consent to contact them. The

diagnosis of cancer was based on the International Classification of Diseases for Oncology-M [81]. After getting physician consent, the questionnaire (see Appendix A) was mailed to the cancer cases (within one to four months of the case's diagnosis). Non-responders were followed-up by telephone.

4.4. Control Recruitment

The Provincial Cancer Registries were also responsible for the recruitment of controls. Controls were selected throughout the 1996 calendar year. A quota of controls to be sampled by age and sex strata was determined for each Province. The methodology of control sampling differed across the Provinces. Specifically, Provincial health insurance registration databases were used in British Columbia, Saskatchewan, Manitoba, Prince Edward Island, and Nova Scotia. Random digit dialling was used in Alberta and Newfoundland. Controls in Ontario were sampled using the Ontario Ministry of Finance Property Assessment Database. In each Province controls were selected by using stratified sampling method so that the age and sex distribution of the entire control series would be similar to that of the entire case series (frequency matching on age and sex). The controls from younger population were over-sampled in order to study types of cancer occurring at younger age. The associations between risk factors and cancer risk would be explored using the same group of controls for different cites of cancer. The same risk factor questionnaire (see Appendix A) was sent to all the identified controls [80].

5. Methods

The analysis of this thesis was based on a sub-set of the National Enhanced Cancer Surveillance System (NECSS) dataset, which is maintained by Cancer Division of Health Canada. The study sample consisted of the subjects diagnosed with bladder cancer and the entire control group selected from seven Provinces (Newfoundland, Prince Edward Island, Nova-Scotia, Manitoba, Saskatchewan, Alberta, and British Columbia). The Provinces of Quebec and New Brunswick did not participate in the NECSS study. The Province of Ontario did not participate in the bladder cancer component of the NECSS because the Provincial Cancer Registry of Ontario was actively involved in another case-control study of bladder cancer.

5.1. Inclusion Criteria

All cases and controls aged 20 years or older at the time of interview were eligible for the inclusion in this study. The study subjects had to be residents in one of seven Canadian Provinces (Newfoundland, Prince Edward Island, Nova-Scotia, Manitoba, Saskatchewan, Alberta, and British Columbia) during the study period, between 1994 and 1997.

The cases had to be diagnosed with bladder cancer as a primary site between January of 1994 and December of 1997, based on the reported topography (C67.0 – C67.9) and histology (transitional cell carcinoma, squamous cell carcinoma, and/or adenocarcinoma), as defined by the International Classification of Diseases for Oncology-M [81]. The controls met the study eligibility criteria if they were free of any type of cancer at their enrolment and had returned a completed questionnaire between 1994 and 1997.

5.2. The Study Sample

The male study sample included 670 bladder cancer cases and 1,642 controls. The female study sample consisted of 359 bladder cancer cases and 1,465 controls.

The overall response rates for male and female bladder cancer cases were 62.0% and 62.4%, respectively. Amongst male and female controls, the corresponding response rates were 64.0% and 70.0%, respectively. Figures 1-4 show sex-specific flow charts depicting the process of recruitment for the cases of bladder cancer and controls separately.

5.3. Power Calculation

The study power was calculated for the samples of men and women separately, using the formula provided by Kelsey [82]. This calculation was based on the comparison of risks between the reference and the highest exposure quartiles of the main exposure variables (recreational physical activity and/or body mass index). Recreational physical activity and body mass index were categorized into quartiles according to their distribution in the controls. Since the comparison was based on the two lowest and highest quartiles, the approximate prevalence of the exposure amongst controls was 0.5 (50%). The number of cases used in the calculation of power was assumed to be about half the total sample of cases, because under the assumption of reasonable OR patterns, the lowest and highest quartiles combined would contain roughly 50% of the total number of cases. Both directions of association (protective effect: OR < 1 and harmful effect: OR > 1) between the main exposure variable and bladder cancer risk were considered. The estimates of

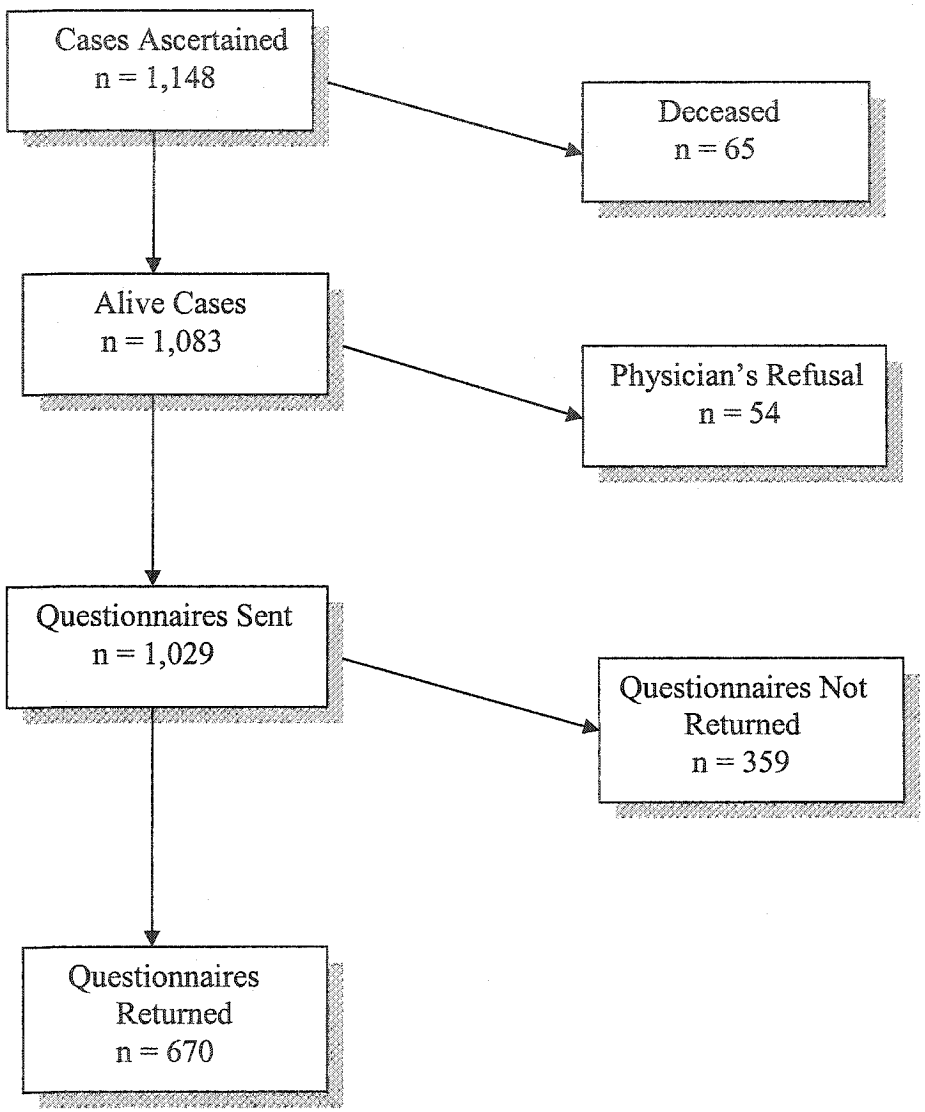
power (expressed in percentages) for the detection of different pre-specified magnitudes of odds ratio (relative risk) are presented in Table 1. For example, in the sample of men, the study would have an adequate power (say at least 80%) of detecting the OR (parameter) of 0.70 or more extreme (or OR of 1.48 or more extreme) if the association of this magnitude (strength) existed in the source population.

Table 1 Power associated with various pre-specified odds ratios/relative risks for the controls' exposure prevalence of 0.5 and the two-sided statistical significance level α = 0.05, NECSS study, Canada, 1994-97

OR/RR*	Power (%)	
	The sample of men 360 cases and 800 controls	The sample of women 180 cases and 740 controls
0.60	98.1	87.9
0.70	80.0	60.0
0.75	61.4	40.0
0.80	46.0	30.5
0.85	23.6	16.0
1.20	23.6	16.0
1.30	46.4	30.5
1.40	70.0	48.4
1.50	87.3	67.0
1.80	99.2	92.1

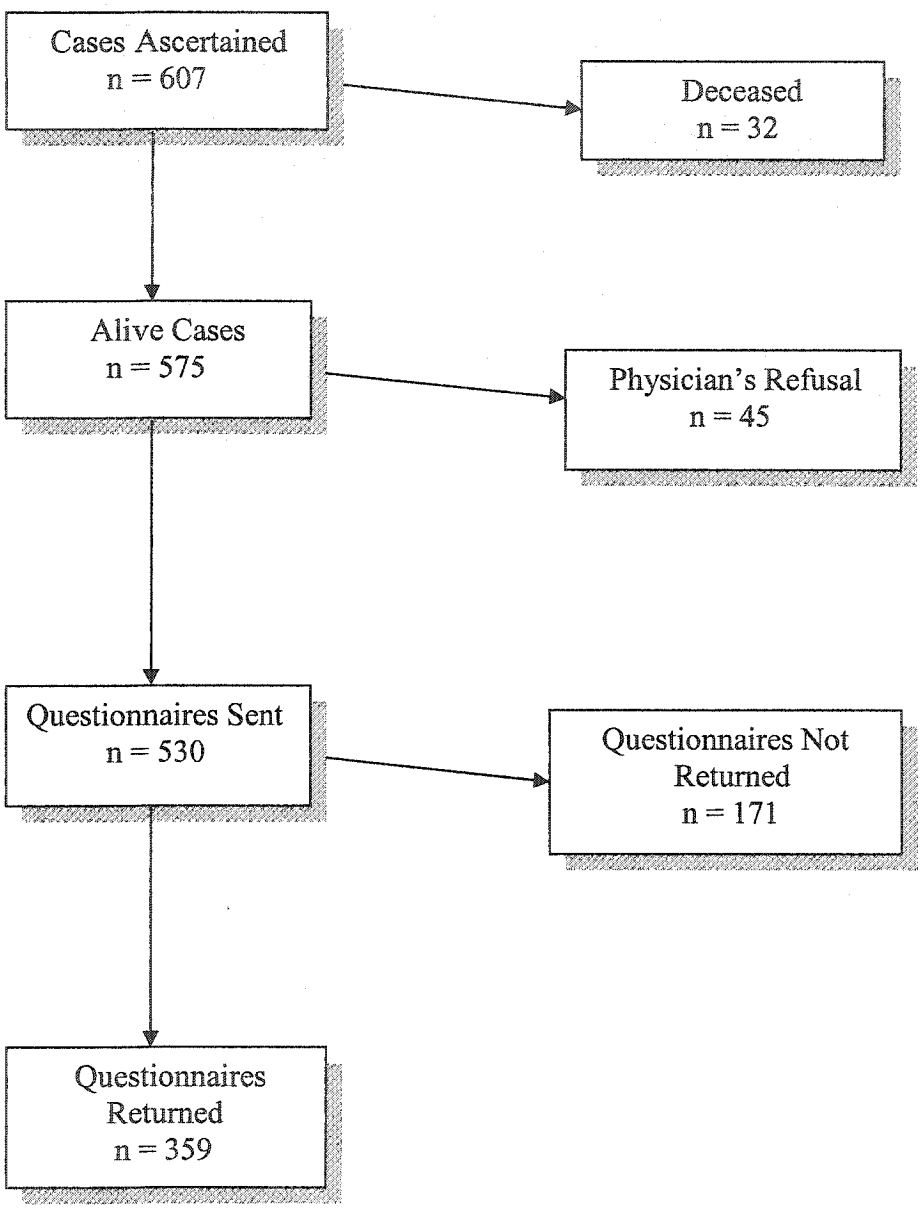
* Odds ratio or relative risk

Figure 1 **Recruitment of Bladder Cancer Cases (Males)**



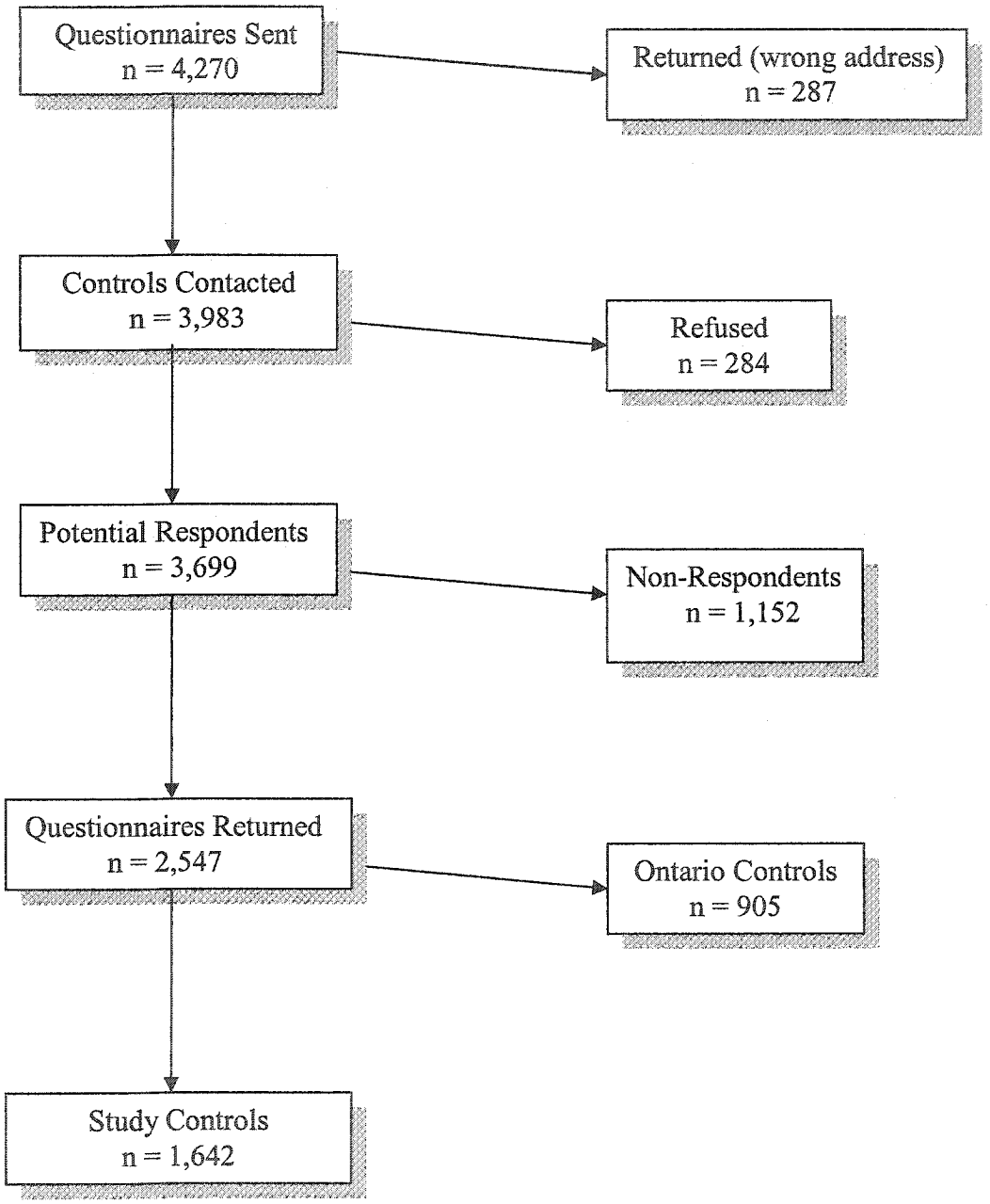
Response rate = 100 x (670/1,083) = 62.0%

Figure 2 **Recruitment of Bladder Cancer Cases (Females)**



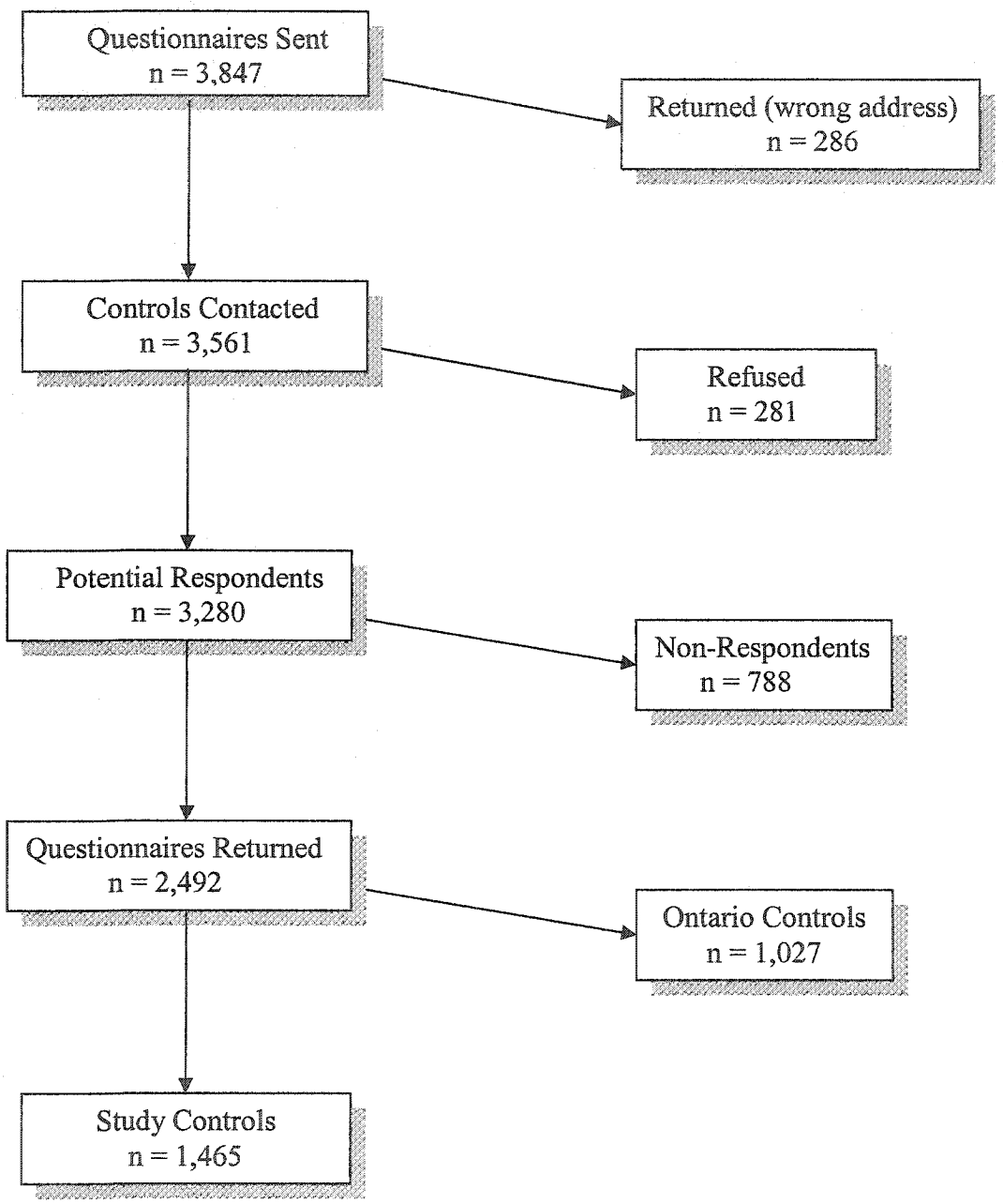
Response rate = 100 x (359/575) = 62.4%

Figure 3 **Recruitment of Controls (Males)**



Response rate = 100 x (2,547/3,983) = 64.0%

Figure 4 **Recruitment of Controls (Females)**



Response rate = 100 x (2,492/3,561) = 70.0%

5.4. Main Exposure Variables

5.4.1. Physical Activity

The completed questionnaire included a separate physical activity section. This section asked a respondent (cases and controls) to provide information on 12 types of recreational activities which had been selected to cover the main types of physical activity in which people engage (walking, jogging/running, gardening/yard work, home exercise, golf, tennis/squash, bowling/curling, swimming/water exercises, skiing/skating, bicycling, social dancing, and other strenuous exercise) (see Appendix A). Specifically, this information was comprised of three components (seasons participated, session frequency, and session duration). Subjects provided this information for each of the 12 types of recreational physical activity. The collected information was cross-sectional and pertained to a point in time approximately two years before a case's diagnosis (or the completion of the questionnaire) or a control's completion of the questionnaire.

The seasonal component of the physical activity questionnaire categorized the study subjects with respect to the number of seasons per year during which they participated in each of the 12 types of activities. This produced a seasonal score which ranged from 0 to 4 (for all 12 types of activities). For example, if a subject was engaged in a certain type of activity (say jogging) only during spring and summer, his or her seasonal score for this particular activity would be 2 per year. In contrast, if this (or any other) subject was not engaged in a specific activity (say swimming) during any of the four seasons, his or her seasonal score for this specific activity would be 0 per year.

Categories for the frequency of each performed session of activity were the following: <1 per month, 1 to 3 per month, 1 to 2 per week, 3 to 6 per week, and every day. Since the first four categories of frequency represented ranges of values rather than discrete numbers and the units were measured in "per week" or "per month," the corresponding categorical designations were converted to a standard number of times per month by using the mid-point of the range (e.g. <1 per month = 0.5, 1 to 3 per month = 2, 1 to 2 per week = 6.5, 3 to 6 per week = 19.4, and every day = 30).

The categories for the duration of each session were: <15 minutes per session, 15 to 30 minutes per session, 31 to 60 minutes per session, and >60 minutes per session. The categorical designations for the duration of each session were standardized to the number of minutes per session by using the mid-point of the range (e.g. <15 minutes per session = 10, 15 to 30 minutes per session = 22.5, 31 to 60 minutes per session = 45.5, and >60 minutes per session = 75).

5.4.2. Body Weight and Height

The questionnaire collected information on self-reported weight (in kilograms) and height (in meters). The reported weight was an estimate (done by a respondent) of a subject's weight two years prior to the diagnosis or the completion of the questionnaire.

5.5. Selection of Covariates

The selection of covariates to be analyzed in this study was guided by the knowledge of their relative importance as risk factors for bladder cancer. This knowledge was based on

the strength of epidemiologic evidence found in the previous literature. The selection process also depended on the availability of covariate data. Certain variables (marital status and income adequacy class) were selected purely for descriptive purposes.

The selected variables (except for physical activity, weight, height, and BMI) for the analysis of this thesis were:

1. Sex (female and male)
2. Age at the interview (years)
3. Province of residence
4. Education (years)
5. Income Adequacy Class (low income, low middle income, upper middle income, and high income)
6. Marital status (married, single, widowed, common law, divorced/separated, and other)
7. Occupation (type of industry/business/service, job title, job duties, job start year, job end year, and Canadian Occupational Standard Codes [83])
8. Cigarette Smoking:
    a) Number of years smoked
    b) Average number of cigarettes smoked per day
    c) Smoking status (2 years before the time of interview): never, former, and current
9. Coffee consumption (number of cups per day: never, < 1, 1, 2–3, and ≥ 4)
10. Alcohol consumption (number of servings per week)
11. Number of times dieted during adult life (never, 1-2, 3-5, 6-8, 9-11, and ≥12)

12. Female reproductive factors:

a) Number of live births

b) Menopausal status at the time of interview (Yes/No)

c) Breast feeding (Never/Ever)

d) Age at first menstruation (years)

e) Age at last menstruation (years)

f) Age at first pregnancy (duration ≥ 5 months) (years)

5.6. Data Cleaning and Missing Values

5.6.1. Physical activity

Some subjects had missing values (blank space in the questionnaire was coded as missing) either for the frequency or duration of an activity session given that they had reported to perform the activity (seasonal score ranged from1 to 4). Instead of excluding the observations with missing values for session frequency and duration, an assumption was made that these activities were not regular/frequent or long enough for a subject to remember and indicate their duration and frequency. Based on the above-mentioned assumption, the missing values for session frequency and duration were substituted by imputing the minimum scores of 0.5 (<1 per month) and 10 (<15 minutes per session), respectively. The number and percentage of observations with missing values for session frequency and duration before the imputation was carried out are presented in the Results section (Table 3). A small proportion of subjects had reported either session frequency or duration for certain types of activities (e.g. walking, bowling/curling, social dancing, and etc…), while their seasonal scores for these activities were set to zero. Usually it is easier

for a subject to report or remember season rather than session duration or frequency of the performed activity. Given this pattern of discrepancy, it was not clear whether the subject had performed the activity or not (see column 'unknown' in Table 3). Therefore, the type-specific (and consequently total) amounts of activity for these subjects were declared as missing.

5.6.2. Occupation

This thesis used the Canadian Occupational Standard Codes [83] to dichotomize the study subjects into exposed (ever held at least one of the well-established 'high risk' occupations for bladder cancer) and non-exposed (never held any of the well-established 'high risk' occupations) groups. Based on the epidemiologic evidence, the following occupations were considered as 'high risk': driver (truck, bus, and etc…), hairdresser, painter, dry cleaner, shoemaker, machinist, carpenter, construction worker, petroleum worker, farmer, and tailor.

The occupational histories were thoroughly examined. These analyses revealed that some subjects (both males and females) had missing (dots or blanks) occupational codes. The patterns of the missing codes were explored in detail to ascertain the reasons for the missing data. For example, a subject could have a missing occupational code for a number of reasons; he or she may have been simply unemployed or retired during the specific period of time for which the occupational code was not assigned and was coded as missing. Another reason could be that even though the subject had worked, his or her

job was not or could not be coded (truly missing), even though some occupational information (job title, job type, job duties, and etc…) for the un-coded job was available.

Taking into consideration the patterns of missing occupational codes, the study subjects initially were grouped according to whether they had ever been employed or not. Within these two groups, several categories were formed based on the pattern of missing (completely or partially) data. Those who had never been employed, had no occupational data, and therefore were assumed to be non-exposed falling into 'non-hazardous' occupational group. The subjects who had ever been employed and missing at least one occupational code, but having some occupational information (job title, job type, job duties, and industry) for the missing code(s), were manually reviewed, their exposure status was ascertained, and assigned by the thesis supervisors and the student independently (females: 28 cases and 48 controls; males: 60 cases and 95 controls). The final exposure status for each subject was reached by consensus. The decision whether the subject was exposed or non-exposed was based on the same principle of the dichotomy of the occupational groups outlined above (if at least 1 occupation 'hazardous' = exposed, otherwise = non-exposed). There was a very small proportion of ever employed study subjects with completely missing occupational data (occupational codes, job titles, job types, job duties, industry, and etc…). These subjects could not be classified ('hazardous' vs. 'non-hazardous) and their exposure status was declared as missing.

5.6.3. Other Variables

Distributions of variables were examined for the detection of implausible (errors or outliers) and missing values. In order to maintain maximum information for descriptive analysis, implausible values were declared as missing, instead of deleting entire observations. Implausible values for height, weight, BMI, reproductive and other variables were explored and sorted out (Table 4). Negative or extremely small values for age (0 year) and height (0 – 1 meter) were treated as missing. BMI values greater than 50 kg/m^2, age at first pregnancy (4 years), and age at last menstruation (3 – 11 or 60 – 77 years) were also declared as missing. A few subjects had missing values for smoking status; the examination of values across other smoking variables for these subjects revealed that they were smokers who quit smoking within two years of the interview date. Since the interest of this study was a smoking status two years before the date of interview, the missing values in this variable for these subjects were re-coded/classified as 'current smoker'. Several women had missing values for pregnancy (ever vs. never), and these values were ascertained by scanning other reproductive variables (i.e. number of live births) for these women. Plausibility of some extremely large numbers of live births for some women was examined by looking at their age at interview.

5.7. Derivation of Total Physical Activity Score and Body Mass Index Calculation

5.7.1. Total Physical Activity Score

Each subject was assigned a total physical activity score (in MET-hours per month) - a sum of the activity levels for each of the 12 activities listed in the questionnaire. Initially,

for each activity, the standardized values of session duration and frequency, and the seasonal score for that specific activity, were multiplied together. Next, this product was multiplied by the metabolic equivalency threshold value (MET) to account for intensity of the activity [84]. The MET value is defined as the ratio of metabolic rate for the specific activity to the resting metabolic rate. One MET is the average seated resting energy cost for an adult and it equals 3.5 ml/kg per minute of oxygen [28; 69; 84-85]. Table 2 displays MET values for the 12 types of activities included in the NECSS study questionnaire. The measurement unit (MET-minutes per year) of the obtained product was then converted to MET-hours per year by dividing this product by 60. And finally, the total score, summarizing the amount of recreational physical activity for all 12 types of activities (measured in MET-hours per month), was obtained by summing the products from all 12 types of activities and dividing this sum by 12 (to convert MET-hours per year to MET-hours per month). Formulae 1.1 and 1.2 show the derivation of the type-specific and the total physical activity scores, respectively:

$$A_{jl} = 3 \times (S_{jl} \times D_{jl} \times F_{jl} \times M_j)/60 \qquad (1.1)$$

$j = \{1, 2, \ldots, 12\}$ Type of activity

$l = \{1, 2, \ldots, n\}$ Subject

A_{jl} – Amount of activity in MET-hours per year for j^{th} type of activity and l^{th} subject

3 – Number of months in season

S_{jl} - Number of seasons per year for j^{th} type of activity and l^{th} subject

D_{jl} – Duration in minutes per session for j^{th} type of activity and l^{th} subject (range mid-point)

F_{jl} - Frequency in times per month for j^{th} type of activity and l^{th} subject (range mid-point)

M_j - MET value for j^{th} type of activity

$$T_1 = \sum_{j=1}^{12} (A_{j1})/12 \qquad (1.2)$$

T_1 - Total score of recreational physical activity in MET-hours per month for 1^{th} subject

Table 2 The 12 types of recreational physical activity and corresponding MET values*, NECSS study, Canada, 1994-97

Type of Activity	MET value
1. Walking	4.0
2. Bowling/curling	3.0
3. Social Dancing	4.5
4. Jogging/running	7.0
5. Gardening/yard work	5.0
6. Home exercise	4.5
7. Golf	4.5
8. Swimming/water exercises	6.0
9. Bicycling	6.0
10. Other strenuous exercise	6.5
11. Tennis/squash	7.0
12. Skiing/skating	7.0

* MET: the ratio of metabolic rate for the specific activity to the resting metabolic rate

A factor 1/12 converts MET-hours per year into MET-hours per month. The total physical activity score (formula 1.2) was used in all subsequent analyses. Specifically, the total score was used to compare the cases and controls with respect to the performed amount of recreational physical activity.

5.7.2. Body Mass Index

Body mass index (measured in kg/m^2), a measure of relative weight, was calculated by dividing self-reported weight in kilograms by the square of self-reported height in meters.

5.8. Derived Covariates

Reproductive factors

The number of live births was used to create a variable for parity. A new reproductive variable measuring the number of years of menstruation was created using the following rules:

For post-menopausal women:

Years menstruated = Age at last menstruation – Age at first menstruation

For pre-menopausal women:

Years menstruated = Age at interview – Age at first menstruation

Based on the hypothesized mechanism, the number of years menstruated (i.e. the reproductive period) could be more important variable than age at first menstruation or age at last menstruation taken separately, since it reflects the duration of exposure to endogenous hormones (estrogen).

Cigarette smoking

Using the following algebraic operation, the number of pack-years of cigarette smoking was calculated, which simultaneously accounts for the amount and the duration of cigarette smoking:

Number of pack-years of cigarette smoking =

= (Number of years smoked) x (Average number of cigarettes smoked per day)/25

5.9. Scale Definition for the Selected Variables

Age

Depending on the purpose of analysis, age was used as either a continuous (in the descriptive part) or categorical variable (in the descriptive and analytical parts). For descriptive purpose, age was categorized in 10-year groups as follows: 20-29, 30-39, 40-49, 50-59, 60-69, and 70 years or older. In the analytical part, the first two age strata (20-29 and 30-39) were collapsed because of extremely small cell numbers.

Province

For descriptive as well as analytical purposes, Province was represented as a nominal variable.

Total physical activity score

In the descriptive analysis, total physical activity score was used as a continuous variable (in MET-hours per month). The rest of the analyses were based on sex-specific quartiles of total physical activity score. The cut-points for these quartiles were based on the score's distribution amongst the study controls. Since men and women perform different

amounts of physical activity, the cut-points and the corresponding quartiles differed between the two groups.

Body mass index

The distribution of body mass index, as a continuous variable, was explored in the descriptive analysis. The cut-points and quartiles of body mass index, used in the analytical part, were determined according to the sex-specific distribution in the study controls. A finer categorization (quintiles and deciles) of body mass index was used to detect any steep increasing or decreasing patterns in risk of bladder cancer, especially for those with BMI $\geq$ 30 kg/m^2.

Occupation

From the study subjects, the NECSS questionnaire collected lifetime employment history, i.e., information on jobs or occupations (including seasonal and part-time work) held for at least 1 year during one's lifetime up to 1997 (most recent job). This information included type of industry/business, main job title, and main job duties for each occupation. The year of start and end for each job was also recorded. Each subject could report a maximum of 12 job/occupational histories ordering them chronologically. Using this information, the NECSS study coded each occupation according to the Canadian Occupational Standard Codes [83].

This thesis used the Canadian Occupational Standard Codes in order to dichotomize the study subjects into hazardous/exposed and non-hazardous/non-exposed occupational groups. For this purpose, using the evidence from epidemiologic literature, the list of

well-established 'high risk' occupations for bladder cancer was created. The following occupations were considered to be hazardous: driver (truck, bus, and etc…), hairdresser, painter, dry cleaner, shoemaker, machinist, carpenter, construction worker, petroleum worker, farmer, and tailor [83]. Afterwards, all the occupational codes corresponding to the above-mentioned occupations were searched and compiled by using the Manual of Canadian Occupational Standard Codes [83]. Based on their occupational codes, the study subjects were then dichotomized into 'hazardous' and 'non-hazardous' occupational groups. Specifically, those who reported to have ever held at least one of the 'high-risk' occupations for at least one year during their lifetime, were classified as exposed ('hazardous' occupational group), otherwise as non-exposed ('non-hazardous' occupational group).

The temporal component (age at starting a hazardous job, hazardous job duration and its chronological place with respect to etiologically relevant time window) of employment history was not incorporated in the occupational variable, because its consideration was beyond the scope of this thesis.

Cigarette smoking

This study used three distinct smoking variables:

1. Ever smoked 100 or more cigarettes (yes/no)
2. Smoking status 2 years before the time of interview (never, former, and current)
3. Pack-years of cigarette smoking

Of the three smoking variables, the number of pack-years of cigarette smoking was chosen to be used in the logistic regression modeling. Since this variable considers not only the amount of smoked cigarettes but also the number of years smoked, it captures more information than one that considers only cross-sectional smoking status (never, former, and current).

Pack-years of cigarette smoking was used both as continuous and categorical variable. The formation of quartiles and cut-points for pack-years of cigarette smoking was based on the sex-specific distribution amongst ever smoker controls. Thus, there were five categories of pack-years: one category for never smokers and four categories or quartiles for ever smokers.

Education

In the descriptive analysis, only continuously distributed variable for education was used (in years). For the rest of the analyses education was categorized as follows: ≤ 9, 10 - 12, and > 12 years.

Coffee and alcohol consumption

The epidemiologic evidence has indicated that low levels of coffee consumption are not associated with an elevated risk of bladder cancer (less than 1 cup per day vs. never) [5; 54-55]. In the study sample, there were very few subjects (about 10-15 %) who had consumed less than 1 cup per day. Given the above-mentioned, the two categories of coffee consumption 'never' and 'less than 1 cup per day' were combined to form a reference

category ('never or less than 1 cup per day'). The risks of bladder cancer associated with upper three categories of coffee consumption (1, 2-3, and ≥ 4 cups per day), relative to the reference category (never drinkers), were similar. Therefore, these categories were combined to form a single category '1 cup per day or more'. Thus, coffee consumption was dichotomized as 'never or less than 1 cup per day' (reference category) versus '1 cup per day or more'.

Alcohol consumption was measured in the number of servings per week. It combined the consumption of beer, wine, and liquor. This variable was categorized according to its sex-specific distribution in the controls. Alcohol consumption was grouped in five categories. The first category included those who never consumed alcohol and the remaining four categories represented quartiles among 'ever drinkers'.

Female reproductive factors

Menopausal status, a dichotomous variable, categorized women as post- and pre-menopausal. Women were classified as postmenopausal if they had stopped menstruating for over a year by the time of interview. Although this variable was used for descriptive purposes, it was not modeled because of strong confounding of age. The examination of 2 by 2 tables (cross-classification of menopausal status by case-control status) stratified across the age groups revealed that none of the cases in the youngest group was post-menopausal and none of the cases and controls in the two oldest groups was pre-menopausal. These non-overlapping distributions of age and menopausal status would

not allow the separation of the effects of age and post-menopausal vs. pre-menopausal status on risk of bladder cancer.

The number of live births was grouped into four categories: 0, 1 – 2, 3 – 4, and ≥ 5. Parity was used as a dichotomous variable (nulliparous vs. parous). Age at first menstruation, age at last menstruation, and age at first pregnancy (duration ≥ 5 months) were dichotomized at 13, 45, and 23 years, respectively (the medians of these distributions in the female controls). The number of years menstruated was dichotomized at 33 years (based on a LOESS smoothed curve of log of odds vs. number of years menstruated).

Income adequacy class

This categorical variable, due to large proportions of missing values for both men (22%) and women (27%), was used only for descriptive purposes. The income adequacy class considered not only the annual household income (in Canadian currency) but also the household size. There were four categories of income adequacy class:

Low income class	$10,000 – 19,999
	$20,000 – 29,000 and household size ≥ 4
Low middle income class	$20,000 – 29,000 and 0≤ household size < 4
	$30,000 – 49,999 and household size ≥ 4

Upper middle income class	$30,000 – 49,999 and 0≤ household size < 4
	$50,000 – 99,999 and household size ≥ 4
High income class	$50,000 – 99,999 and 0≤ household size < 4
	> $100,000

Number of times dieted during adult life

This variable was used only for descriptive purposes and it was grouped into the following 6 categories: never, 1 – 2, 3 – 5, 6 – 8, 9 – 11, and ≥ 12.

5.10. Data Analysis

Given the distinct hormonal milieu in men and women, hormonal mediation of the effects of physical activity and body fat mass (relative body weight) on bladder cancer risk may differ across the two groups.

Female reproductive factor(s), as surrogate measure(s) of exposure to endogenous estrogens, may be related to risk of female bladder cancer independently of physical activity and/or relative body weight.

All analyses were performed separately for men and women.

5.10.1. Descriptive Statistics

The descriptive analysis was performed for categorical as well as continuous variables. The distributions of all variables across cases and controls were generated and examined.

The frequency distribution of cases and controls for categorical variables were expressed in numbers and percentages. The distributions of continuous variables were examined by calculating the statistics of central tendency and spread (means, medians, ranges, standard deviations, 25^{th} and 75^{th} percentiles). The shapes of the distributions of various continuous variables were examined through plotting histograms. The shape of the distributions aided in the decision as to which summary statistics to use for more appropriate characterization of the distribution. Besides revealing the basic descriptive findings, the descriptive analysis was also used to explore patterns of missing data and to detect outliers or implausible values for the purposes of data cleaning.

5.10.2. Distributions of Total Physical Activity Score and Body Mass Index in Controls

The previous epidemiologic research has indicated that the prevalence of exposure differs across certain subgroups. For example, it has been often shown that men tend to drink more alcohol or smoke more cigarettes than women, physical activity levels are higher in younger than older population, subjects with higher socioeconomic status are more physically active (leisure activity) than those with lower socioeconomic status, or that subjects with lower socioeconomic status tend to be more obese than those with higher socioeconomic status [28].

However, gross violations in the ways the study controls are selected and their exposure is measured may produce unexpected patterns of the exposure distribution amongst the subsets of controls. For example, if the problems related to the controls' selection or

exposure measurement were more severe (or differential) for some subsets than others, one would observe unusual patterns of exposure prevalence across these subgroups of the control series [86]. Confirming the presence of the expected distributions of exposure amongst controls might be indicative that there were no large violations in the selection of controls and the measurement of exposure.

This analysis examined the presence/absence of expected distributional patterns (based on the external knowledge) in sub-groups of the study controls. Specifically, the distributions (box-plots) of total physical activity score and body mass index across different subsets of the controls were generated (for men and women separately). These subsets were categories of age, cigarette smoking, education, income adequacy, the number of times dieted during adult life, and the number of live births. Appendices B1 and B2 display the box-plots (see also the corresponding Tables 8-9 and 18-19).

5.10.3. Model-Building Strategies

The case-control data was analyzed using unconditional logistic regression, where the model parameters were estimated by the maximum likelihood methods [87]. The associations between the independent variable(s) and risk of bladder cancer were examined by generating odds ratios (ORs) and 95 percent confidence intervals (95% CIs), adjusted for age, Province, and other covariates. The model-building strategy was based on the forward stepwise selection of independent variables.

Logit plots (log odds against mid-points of quartiles/quintiles) were constructed to examine the patterns of relationships between independent variables and log odds of bladder cancer. This procedure helped to determine the optimal scale for the modeled variables.

Bivariable method for the selection of variables based on crude models was avoided, since it might wrongly reject potentially important variable due to negative confounding [88]. Therefore, the associations between each independent variable (levels of physical activity, body mass index, and other covariates) and risk of bladder cancer, obtained from the initial logistic models, were adjusted at least for age and Province. At this stage, the cut-point of statistical significance for retaining an independent variable in the model was set at $\alpha = 0.25$.

Using multivariable logistic regression analysis, the effects of physical activity (main exposure variable), body mass index (main exposure variable), and other independent variables were examined for potential confounding due to another covariate by comparing the magnitudes of OR estimates obtained from the models with and without this covariate. If the magnitude of OR changed by more than 10%, then the covariate was retained for further modeling analyses [87].

The statistical significance of each parameter was assessed by using 95% CI and/or the Wald test. A likelihood ratio test was used to assess the statistical significance of the model parameter(s) taken as a group.

The tests of trend were performed for ordinal categorical variables to determine whether the association between the categories (relative to the reference category) and risk of bladder cancer conformed linear gradient (dose-response relationship). Linear trends for ordinal categorized data were tested by modeling (multivariable logistic regression) the median values of quartiles as a continuous variable. The statistical significance of linear trend was assessed by Wald Chi-square test statistic (two-sided P-value). The level of statistical significance for this procedure was set at $\alpha = 0.05$.

All main exposure variables were retained in the final models regardless of their overall importance. The decision to retain a covariate in a final model was based not only on the 95% CI and the magnitude of the P value, but also on its clinical relevance. The decision depended also on how strongly the covariate confounded the associations between main exposure variables and the outcome. Efforts were taken to obtain parsimonious final models. Thus, the automatic decisions based on the simple dichotomy determined by the arbitrary $\alpha = 0.05$ level, were avoided. For example, if a covariate was not statistically significant at the $\alpha = 0.05$ level, but it showed a strong confounding effect and was a well-known risk factor, it was kept in the model. Whereas, if a well-known risk factor did not affect the effect estimates of main exposure variables and its statistical significance was greater than $\alpha = 0.25$, it was not retained in the final model. Those variables that were initially screened out as not important in terms of their contribution to the model were re-evaluated also in the final models.

The previous research of female cancer has shown that risk factors for breast and ovarian cancers operate differently in pre- and post-menopausal women [89-91]. Therefore, the possibility of potential effect modification by menopausal status was explored by building logistic models based on the samples of pre- and post-menopausal women. Given a reduced risk of bladder cancer found in this study amongst parous women (having at least one live birth) relative to nulliparous women, the association between the number of live births and bladder cancer risk was examined separately in the sample of parous women. This analysis was intended to explore whether having a greater number of live births was related to any additional decrease in risk of bladder cancer.

The fit of the final model was assessed using Hosmer-Lemeshow (Chi-square) goodness of fit test.

5.10.4. Software Used for the Analyses

All calculations for this analysis were performed using Statistical Analysis Systems (SAS, version 8.2) [92].

6. Results

6.1. Data Cleaning and Missing Values

6.1.1. Recreational Physical Activity

The amount of initially missing data for session frequency and duration of the 12 types of recreational physical activity among those who reported any activity varied between 0.1 and 6.7 % (Table 3). As explained in the Methods section, about 5% of the total sample was imputed by the minimum values for frequency and duration components. After these imputations, the total physical activity score was computed. The remaining missing values were for subjects who had reported either the session frequency or duration for certain specific activities, while the seasonal scores for these activities were set to zero, meaning that they had not performed this type of activity. Because of this discrepancy, it was not known whether these subjects had performed the activity or not (see the 'unknown' column in Table 3). Consequently, the amount of type-specific activities and total physical activity scores for these subjects (men: 34 cases and 70 controls; women: 15 cases and 48 controls) could not be computed and were declared as missing (see Tables 5 and 13).

6.1.2. Covariates

The descriptive examination of the variables revealed some implausible values, but only amongst the study controls. The extreme values were observed for the following variables: age, height, body mass index, the number of live births, age at first menstruation, age at last menstruation, age at first pregnancy (duration ≥ 5 months), and years menstruated. The data are presented in Table 4.

Table 3 Proportions of observations with initially missing values for session frequency and duration across 12 types of physical activity, NECSS study, Canada, 1994-97

Type of activity	Cases (1,029 men and women)					Controls (3,107 men and women)				
				Missing					Missing	
	No*	Yes#	Unknown§	Frequency Number (%)¶	Duration	No	Yes	Unknown	Frequency Number (%)	Duration
Walking	284	735	10	14 (1.3)	21 (2.0)	784	2,299	24	41 (1.3)	73 (2.4)
Jogging/running	975	50	4	2 (0.2)	4 (0.4)	2,814	279	14	8 (0.2)	11 (0.3)
Gardening	241	782	6	44 (4.3)	69 (6.7)	759	2,331	17	97 (3.1)	203 (6.5)
Home exercise	819	201	9	6 (0.6)	13 (1.3)	2,187	891	29	19 (0.6)	45 (1.4)
Golf	798	228	3	3 (0.3)	20 (1.9)	2,486	612	9	15 (0.5)	46 (1.5)
Tennis	1,000	29	0	0 (0.0)	2 (0.2)	2,916	184	7	2 (0.1)	5 (0.2)
Bowling	849	176	4	5 (0.5)	14 (1.4)	2,590	497	20	7 (0.2)	38 (1.2)
Swimming	776	249	4	7 (0.7)	41 (4.0)	2,249	846	12	31 (1.0)	86 (2.7)
Skiing/skating	899	129	1	7 (0.7)	11 (1.0)	2,504	594	9	16 (0.4)	25 (0.8)
Bicycling	790	238	1	7 (0.7)	20 (1.9)	2,289	816	11	26 (0.8)	60 (1.9)
Social dancing	719	287	23	5 (0.5)	27 (2.6)	2,144	906	57	21(0.6)	124 (4.0)
Other str. Exer†.	870	154	5	6 (0.6)	15 (1.5)	2,527	563	17	27 (0.8)	33 (1.0)

* Subjects who had not performed the activity

\# Subjects who had performed the activity

§ Unknown whether the subjects had performed the activity or not

¶ For comparison across the 12 types of activities, the percentages were calculated using the constant denominators of 1,029 cases and 3,107 controls

† Other strenuous exercise

Table 4 Variables and corresponding implausible values, NECSS study, Canada, 1994-97

Covariate	Actual Value	N*
Age (years)	{-11, 0}	3
Height (meters)	{0, 1.0, 3.3}	8
Body mass index (kg/m^2)	{51, 53, 54, 57, 59, 68, 85, 93}	14
Number of live births (count)	{22, 25, 31}	3
Age at first pregnancy (years)	{4}	1
Age at last menstruation (years)	{3, 4, 5, 11} and {59, 60, 77}	10
Age at first menstruation (years)	{1}	1
Number of years menstruated (years)	{51, 55, 56, 57, 67}	7

* Number of observations with an outlier value

The declaration of these implausible values as missing did not influence prior distributions of the above-mentioned variables. The final amount of missing data (originally plus after declaring implausible values as missing) for each of the covariates is presented in the descriptive tables of Results section.

6.2. Analyses of the Sample of Men

6.2.1. Description of the Study Sample

The distribution of main exposure variables (total physical activity score and body mass index) as well as other socio-demographic and lifestyle characteristics in 670 male cases and 1,642 male controls, are presented in Tables 5 – 7.

Table 5 Selected socio-demographic, anthropometrical, and life-style characteristics of male cases and controls, NECSS study, Canada, 1994-97

Characteristic	Cases (n = 670)			Controls (n = 1,642)		
	Range	Mean (SD)*	Missing #	Range	Mean (SD)	Missing
Age (years)	29 – 75	63.2 (8.9)	1 (0.2)	20 – 76	57.9 (14.6)	7 (0.4)
BMI (kg/m^2) §	15.7 – 44.7	26.7 (3.9)	1 (0.2)	14.6 – 45.6	26.2 (3.8)	16 (1.0)
Education (years)	3 – 32	11.4 (3.6)	10 (1.5)	1 – 27	11.9 (3.9)	25 (1.5)
Total physical activity score†	0 – 795	106 (105.5)	34 (5.0)	0 – 995	103 (104.4)	70 (4.3)
		Median (25% quartile, 75% quartile) value				
Total physical activity score	0 – 795	78.5 (25.0, 153.9)	34 (5.0)	0 – 995	73.6 (24.2, 150.0)	70 (4.3)
Smoking (pack-years)	0 - 112	24.0 (12.0, 36.0)	12 (1.8)	0 - 141	12.0 (0.0, 28.0)	31 (1.9)
Alcohol consumption ‡	0 - 105	3.0 (0.0, 8.5)	0 (0.0)	0 - 126	2.0 (0.0, 7.5)	2 (0.1)

* Standard deviation

\# Number and percentage (%) of missing values

§ Body mass index

† Number of MET-hours per month

‡ Combined beer, wine, and liquor consumption in number of servings per week

Because some variables displayed a strongly skewed distribution, they are presented in terms of their medians, 25th and 75th percentiles. The mean age of cases was 63.2 years and that of controls was 57.9 years. This difference in age partially reflects the fact that controls had been over-sampled from the younger age strata (see Table 6). Based on the means displayed in Table 5, the male cases and controls were similar with respect to their body mass index, education, and total physical activity scores. The cases had a slightly higher median total physical activity score than the controls (78.5 vs. 73.6). Similarly, alcohol consumption (median = 3.0) amongst cases was higher than that amongst controls (median = 2.0). The cases had twice the median number of pack-years of smoking than the controls (24.0 vs. 12.0).

According to Table 6, about 60% of the cases and 50% of the controls were from Alberta and British Columbia. About 37% of the controls had more than 12 years of education, as opposed to 28.8% of the cases. The proportions of those who had ever held any hazardous occupation amongst cases and controls were 44.5% and 35.7%, respectively.

Table 7 displays the distribution of cases and controls with respect to smoking, alcohol, and coffee consumption categories. Based on this table, cases were more likely to be ever smokers than controls (88.4% vs. 74.7%). About 30% of cases were heavy smokers (pack-years≥ 35) as compared to only 18% of heavy smokers (pack-years≥ 35) amongst controls. Those who had at least 1 cup of coffee per day were more prevalent in cases than in controls (81.8% vs. 69.0%).

Table 6 Frequencies of male cases and controls according to selected characteristics, NECSS study, Canada, 1994-97

Characteristic	Cases (n = 670) Number (%)[§]	Controls (n = 1,642) Number (%)
10-year age groups		
20 – 29	1 (0.1)	99 (6.1)
30 – 39	10 (1.5)	176 (10.7)
40 – 49	52 (7.8)	137 (8.4)
50 – 59	126 (18.8)	239 (14.6)
60 – 69	283 (42.3)	581 (35.5)
70 or older	197 (29.4)	403 (24.6)
Missing values	1 (0.2)	7 (0.4)
Province		
Newfoundland	42 (6.3)	124 (7.5)
Prince Edward Island	15 (2.2)	82 (5.0)
Nova Scotia	64 (9.6)	337 (20.5)
Manitoba	90 (13.4)	156 (9.5)
Saskatchewan	64 (9.5)	148 (9.0)
Alberta	199 (29.7)	327 (19.9)
British Columbia	196 (29.2)	468 (28.5)
Education (in years)		
≤ 9	201 (30.4)	426 (26.3)
10 – 12	269 (40.8)	589 (36.4)
>12	190 (28.8)	602 (37.2)
Missing values	10 (1.5)	25 (1.5)
Marital status		
Common law	20 (3.0)	51 (3.1)
Divorced/Separated	41 (6.2)	99 (6.0)
Married	552 (82.8)	1,262 (77.0)
Single	32 (4.8)	168 (10.3)
Widowed	21 (3.1)	55 (3.4)
Other	-	3 (0.2)
Missing values	4 (0.6)	4 (0.2)
Occupational group		
Non-hazardous	370 (55.5)	1,051 (64.3)
Hazardous	297 (44.5)	583 (35.7)
Missing values	3 (0.4)	8 (0.5)

[§] Due to rounding errors, percentages (denominators exclude missing values) of the observed values may not add up to 100

Table 7 Frequencies of male cases and controls with respect to the categories of cigarette smoking, alcohol, and coffee consumption, NECSS study, Canada, 1994-97

Characteristic	Cases (n = 670) Number (%)[§]	Controls (n = 1,642) Number (%)
Smoked 100+ cigarettes		
Yes	592 (88.4)	1,226 (74.7)
No	78 (11.6)	416 (25.3)
Smoking status[*]		
Never smoked	78 (11.6)	416 (25.3)
Ex-smoker	368 (54.9)	815 (49.6)
Current smoker	224 (33.4)	411 (25.0)
Pack-years of smoking[**]		
0	78 (11.8)	416 (25.8)
1 – 8	61 (9.3)	302 (18.7)
9 – 19	134 (20.4)	296 (18.4)
20 – 34	192 (29.2)	303 (18.8)
≥35	193 (29.3)	294 (18.2)
Missing values	12 (1.8)	31 (1.9)
Coffee consumption (per day)		
Never or less than 1 cup	121 (18.2)	500 (31.0)
1 cup or more	543 (81.8)	1,111 (69.0)
Missing values	6 (0.9)	31 (1.9)
Alcohol consumption (servings per week) [***]		
0	195 (29.1)	506 (30.8)
>0 to ≤ 1.5	115 (17.1)	288 (17.6)
>1.5 to ≤ 5.5	114 (17.0)	282 (17.2)
>5.5 to ≤11.0	109 (16.3)	289 (17.6)
>11.0	137 (20.4)	275 (16.7)
Missing values	0 (0.0)	2 (0.1)

[§] Due to rounding errors, percentages (denominators exclude missing values) of the observed values may not add up to 100

[*] Two years before the interview

[**] Quartiles based on the distribution in smoker male controls

[***] Combined consumption: beer, wine, and liquor. Quartiles based on the distribution of alcohol consumption in male drinker controls

6.2.2. Total Physical Activity Score and Body Mass Index across the Subsets of Male Controls

The distribution of total physical activity score across different subsets of the male controls presented in Table 8 and the corresponding box-plots (see Appendix B1) indicate that more physically active men tended to be younger and to have higher levels of education and income. Physical activity levels were lower in heavy smokers (≥ 35 pack-years) than in never smokers (means: 78.3 vs. 110.8 MET-hours per month).

Table 8 Summary statistics for total physical activity score across age, smoking, educational, and income adequacy groups in male controls, NECSS study, Canada, 1994-97

Group	Mean	25%	Median	75%	N*
Age (years)					
< 50	117.7	33.4	87.6	171.1	397
50 – 70	98.9	23.8	74.7	145.1	854
> 70	84.6	16.9	59.2	128.1	306
Smoking (pack-years)					
0	110.8	30.7	74.4	156.5	400
1 – 8	117.4	34.4	91.6	171.7	286
9 – 19	105.5	24.5	85.1	154.2	284
20 – 34	85.0	18.8	67.1	123.7	289
≥ 35	78.3	9.8	53.7	121.2	277
Education (years)					
0 – 9	71.0	8.6	44.1	98.0	175
10 – 12	105.7	30.1	72.7	160.4	275
≥ 13	121.1	33.4	93.8	172.2	330
Income adequacy					
Low	88.8	13.1	60.3	137.3	152
Lower middle	96.8	27.8	68.1	150.8	170
Upper middle	117.2	29.1	79.1	170.7	221
High	117.7	45.0	97.9	171.7	109

* Sample size

The subjects in 1-8 pack-years of smoking category tended to have the highest physical activity level. Most of these subjects may have been ex-smokers who increased their activity levels after quitting smoking. The distribution of body mass index across different subsets of the male controls, displayed in Table 9 and the corresponding box-plots (see Appendix B1), indicate that men with lower educational level tended to have higher relative weight and to diet more frequently during their adult lives. There was not any discernible trend in the distribution of body mass index across the four income adequacy groups. In summary, the distributions of total physical activity score and body mass index across the control subsets conformed to the expected patterns.

Table 9 Summary statistics for body mass index across diet, income adequacy, and educational groups in male controls, NECSS study, Canada, 1994-97

Group	Mean	25%	Median	75%	N*
Number of times dieted ‡					
Never	25.0	23.0	24.8	26.6	1,025
1 – 2	27.5	25.1	27.2	29.3	323
3 – 5	29.0	26.3	28.1	31.1	160
≥ 6	29.3	25.8	29.7	32.3	94
Income adequacy					
Low	26.0	23.0	25.6	28.4	280
Lower middle	26.1	23.8	25.7	28.0	313
Upper middle	26.0	23.2	25.6	28.1	421
High	26.1	23.6	25.6	27.8	273
Education (years)					
0 – 9	26.9	24.2	26.4	29.1	423
10 – 12	26.4	23.8	25.8	28.7	584
≥ 13	25.5	23.2	25.1	27.4	597

* Sample size

‡ The number of times dieted during adult life

6.2.3. Model-Based Analysis

Estimates of age- and Province-adjusted OR of male bladder cancer for levels of recreational physical activity (the quartiles of total physical activity score), body mass index, and other covariates are presented in Tables 10-11. The ORs and 95% CIs suggest no association between the levels of recreational physical activity and risk of bladder cancer in men. The age and Province-adjusted risks in the two highest quartiles of body mass index were increased compared to the lowest quartile (OR = 1.38, 95% CI: 1.04, 1.83 and OR = 1.27, 95% CI: 0.97, 1.67). The test for trend for the association between the levels of body mass index and risk of bladder cancer was statistically significant at α = 0.05 level (P value = 0.03). Men who had ever been employed in any hazardous occupations for at least one year were at increased risk of having bladder cancer (age- and Province-adjusted OR = 1.33, 95% CI: 1.09, 1.61). An inverse but marginally statistically significant association between the years of education and risk of bladder cancer was observed only for those with at least 12 years of education (age- and Province-adjusted OR = 0.77, 95% CI: 0.60, 1.00).

The estimates of relationship between smoking, coffee, alcohol consumption and bladder cancer risk are displayed in Table 11. As was expected, based on results of the previous research, cigarette smoking was strongly associated with bladder cancer risk. Ever smokers had a statistically significantly increased risk of bladder cancer compared with never smokers (age- and Province-adjusted OR = 2.21, 95% CI: 1.69, 2.90).

Table 10 Age* - and Province-adjusted odds ratios for recreational physical activity, body mass index, and other characteristics in male cases (n=670) and controls (n=1,642), NECSS study, Canada, 1994-97

Characteristic	OR	95% CI †	Number of subjects** Cases/Controls
Total physical activity score (MET-hours/month) ‡			635/1,566
≥0 to ≤ 24.2	1.00¶		
> 24.2 to ≤73.6	1.03	0.78, 1.36	
>73.6 to ≤150.0	1.03	0.78, 1.35	
>150.0	1.08	0.82, 1.43	
P value for trend	0.60		
Body mass index (kg/m^2) §			669/1,622
< 23.6	1.00¶		
≥23.6 to < 25.7	1.00	0.75, 1.33	
≥25.7 to < 28.0	1.38	1.04, 1.83	
≥28.0	1.27	0.97, 1.67	
P value for trend	0.03		
Occupational group			666/1,627
Non-hazardous	1.00¶		
Hazardous	1.33	1.09, 1.61	
Education (in years)			660/1,623
≤ 9	1.00¶		
10 – 12	1.03	0.81, 1.30	
>12	0.77	0.60, 1.00	
P value for trend	0.03		

* Age categorized as 20-39, 40-49, 50-59, 60-69, and 70 years or older

** Subjects with missing values for the variable were excluded from the analysis

† 95 percent confidence interval

‡ Quartiles based on the distribution of total physical activity score in male controls

§ Quartiles based on the distribution of body mass index in male controls

¶ Reference category

Table 11 Age* - and Province-adjusted odds ratios for smoking, coffee and alcohol consumption groups in male cases (n=670) and controls (n=1,642), NECSS study, Canada, 1994-97

Characteristic	OR	95% CI †	Number of subjects** Cases/Controls
Smoked 100+ cigarettes			669/1,635
No	1.00¶		
Yes	2.21	1.69, 2.90	
*Smoking status****			669/1,635
Never smoked	1.00¶		
Ex-smoker	1.88	1.42, 2.50	
Current smoker	2.96	2.18, 4.02	
Pack-years of smoking ‡			658/1,607
0	1.00¶		
1 – 8	1.08	0.74, 1.60	
9 – 19	2.06	1.48, 2.87	
20 - 34	2.78	2.03, 3.81	
≥35	2.95	2.14, 4.06	
P value for trend	< 0.01		
Coffee consumption (per day)			663/1,605
Never or less than 1 cup	1.00¶		
1 cup or more	1.56	1.23, 1.98	
Alcohol consumption (servings per week) §			669/1,633
0	1.00¶		
>0 to ≤1.5	1.10	0.83, 1.47	
>1.5 to ≤5.5	1.09	0.81, 1.44	
>5.5 to ≤11.0	1.03	0.77, 1.37	
>11.0	1.28	0.97, 1.70	
P value for trend	0.10		

* Age categorized as 20-39, 40-49, 50-59, 60-69, and 70 years or older

** Subjects with missing values for the variable were excluded from the analysis

*** Two years before the interview

† 95 percent confidence interval

‡ Quartiles based on the distribution of pack-years in male smoker controls

§ Combined consumption: beer, wine, and liquor. Quartiles based on the distribution of alcohol consumption in male drinker controls

¶ Reference category

Former and especially current cigarette smokers were at higher risk of developing bladder cancer than never smokers (current vs. never: age- and Province-adjusted OR = 2.96, 95% CI: 2.18, 4.02). The number of pack-years of cigarette smoking was also strongly associated with risk of bladder cancer. The smokers in the highest quartile (≥ 35 pack-years) had almost three times the risk of never smokers (age- and Province-adjusted OR = 2.95, 95% CI: 2.14, 4.06). Moreover, there was a strong statistically significant linear trend in the risk gradient associated with the number of pack-years of cigarette smoking (P value < 0.01). The association between higher coffee consumption and risk of bladder cancer adjusted for age and Province was positive and statistically significant at $\alpha = 0.05$ (OR = 1.56, 95% CI: 1.23, 1.98). As for the risk associated with alcohol consumption, only the men in the highest quartile of alcohol consumption (>11 servings per week) had a slightly increased risk of bladder cancer compared to those who had never used alcohol (age- and Province-adjusted OR = 1.28, 95% CI: 0.97, 1.70). The test for trend for the association between the levels of alcohol consumption and bladder cancer risk was not statistically significant at $\alpha = 0.05$ level (P value = 0.10).

The relationship between the levels of recreational physical activity and risk of bladder cancer obtained from the final multivariable logistic regression model are presented in Table 12. The odds of bladder cancer for the two total physical activity score quartiles > 24.2 to ≤ 73.6 and > 73.6 to ≤ 150.0 MET-hours/month were similar to that for the reference category of ≥ 0 to ≤24.2 MET-hours/month (OR = 1.05, 95% CI: 0.79, 1.40 and OR = 1.04, 95% CI: 0.78, 1.38). Although the odds of bladder cancer were slightly elevated for the highest quartile of physical activity score, the effect estimate was not statistically significant (OR = 1.24, 95% CI: 0.93, 1.66).

Table 12 Multivariable odds ratios for male bladder cancer based on 617 cases and 1,499 controls, NECSS, Canada, 1994-97

Characteristic	OR[#]	95% CI [†]
Total physical activity score (MET-hours/month) [‡]		
≥0 to ≤ 24.2	1.00[¶]	
> 24.2 to ≤ 73.6	1.05	0.79, 1.40
> 73.6 to ≤ 150.0	1.04	0.78, 1.38
> 150.0	1.24	0.93, 1.66
P value for trend	0.13	
Body mass index (kg/m^2)[*]		
< 23.6	1.00[¶]	
≥ 23.6 to < 25.7	1.05	0.78, 1.42
≥ 25.7 to < 28.0	1.49	1.10, 2.01
≥ 28.0	1.29	0.96, 1.72
P value for trend	0.10	
Occupational Group		
Non-hazardous	1.00[¶]	
Hazardous	1.22	1.00, 1.50
Coffee consumption (per day)		
Never or less than 1 cup	1.00[¶]	
1 cup or more	1.30	1.00, 1.68
Pack-years of smoking[§]		
0	1.00[¶]	
1 – 8	1.04	0.70, 1.54
9 – 19	1.86	1.32, 2.62
20 – 34	2.61	1.88, 3.63
≥ 35	2.67	1.91, 3.75
P value for trend	< 0.01	

[#] Odds ratio adjusted for all the variables listed in this table (except for the variable under consideration), age (20-39, 40-49, 50-59, 60-69, and 70 years or older), and Province

[†] 95 percent confidence interval

[‡] Quartiles based on the distribution of total physical activity score in male controls

[*] Quartiles based on the distribution of body mass index in male controls

[§] Quartiles based on the distribution of pack-years in male smoker controls

[¶] Reference category

The P value for trend (P value = 0.13) also was not supportive of an increasing linear gradient in the risk across the levels of physical activity. The risks associated with two upper body mass index quartiles (≥ 25.7 to < 28.0 and ≥ 28.0 kg/m^2) were somewhat increased relative to the reference category, but both estimates were marginally statistically significant (OR = 1.49, 95% CI: 1.10, 2.01 and OR = 1.29, 95% CI: 0.96, 1.72). The test for trend for the association between the levels of body mass index and bladder cancer risk yielded a P value of 0.10.

Those who had ever been occupied in 'high risk occupations' were at higher odds of having bladder cancer, although the estimate bordered the margin of statistical significance (OR = 1.22, 95% CI: 1.00, 1.50). The estimate of OR and 95% CI for the association between coffee consumption (1 cup or more vs. never or less than 1 cup a day) and bladder cancer indicate a possible increase in the risk of bladder cancer for those drinking 1 cup or more a day. In the final model, the previously observed strongly positive relationship between the number of pack-years of cigarette smoking and bladder cancer risk persisted, i.e., large ORs (ranging from 1.86 to 2.67) were accompanied by strong statistical evidence against the null hypothesis (95% CIs and P value for trend).

The Hosmer-Lemeshow goodness-of-fit test for the final multivariable model yielded a Chi-square (8 degrees of freedom) of 12.0 (P value = 0.14), indicating a satisfactory fit of the model.

6.3. Analyses of the Sample of Women

6.3.1. Description of the Study Sample

The distributions of total physical activity score, body mass index, and other socio-demographic/lifestyle characteristics in 359 female cases of bladder cancer and 1,465 female controls, are presented in Tables 13 – 15. A skewed distribution of certain variables is presented in terms of range, median, 25th and 75th percentiles. Mean age values for female cases and controls were 61.1 and 56.2 years, respectively. Cases and controls had similar mean values of body mass index and education (Table 13). Both, the mean and the median of total physical activity score were slightly higher in controls than in cases (means: 82.3 vs. 76.4 MET-hours/month; medians: 60.2 vs. 50.5 MET-hours/month). The median number of pack-years of smoking was higher in cases than in controls (13.0 vs. 1.0).

The frequencies of cases and controls in the 10-year age groups are displayed in Table 14. Women controls, compared to cases, included a higher proportion of younger subjects. The majority of the female cases (45%) and controls (48%) were from British Columbia and Alberta. Controls were somewhat more educated than cases. The proportions of women who had ever held any hazardous occupation(s) for bladder cancer across cases and controls were 19.3% and 12.7%, respectively.

Frequencies of cases and controls with respect to smoking, coffee and alcohol consumption categories are shown in Table 15. About 72.0% of the cases as opposed to

Table 13 Selected socio-demographic, anthropometrical, and life-style characteristics of female cases and controls

Characteristic	Cases (n = 359)			Controls (n = 1,465)		
	Range	Mean (SD)*	Missing #	Range	Mean (SD)	Missing
Age (years)	22 – 75	61.1 (10.8)	0 (0.0)	20 – 76	56.2 (12.2)	2 (0.1)
BMI (kg/m^2) §	14.7 – 46.3	25.6 (5.0)	1 (0.3)	15.4 – 44.5	25.3 (4.8)	11 (0.7)
Education (years)	2 – 23	11.5 (2.8)	3 (0.8)	2 – 37	12.0 (3.1)	15 (1.0)
Total physical activity score†	0 – 449	76.4 (80.4)	15 (4.2)	0 – 529	82.3 (77.3)	48 (3.3)
		Median (25% quartile, 75% quartile) value				
Total physical activity score	0 – 449	50.5 (19.6, 106.8)	15 (4.2)	0 – 529	60.2 (22.5, 121.8)	48 (3.3)
Smoking (pack-years)	0 - 98	13.0 (0.0, 26.0)	6 (1.7)	0 - 80	1.0 (0.0, 12.0)	22 (1.5)
Alcohol consumption ‡	0 - 63	0.0 (0.0, 3.0)	0 (0.0)	0 – 43	0.5 (0.0, 2.0)	2 (0.1)

* Standard deviation
\# Number and percentage (%) of missing values
§ Body mass index
† Number of MET-hours per month
‡ Combined beer, wine, and liquor consumption in number of servings per week

Table 14 Frequencies of female cases and controls according to selected characteristics

Characteristic	Cases (n = 359) Number (%)[§]	Controls (n = 1,465) Number (%)
10-year age groups		
20 – 29	4 (1.1)	42 (2.9)
30 – 39	11 (3.1)	97 (6.6)
40 – 49	39 (10.8)	287 (19.6)
50 – 59	77 (21.4)	366(25.0)
60 – 69	136 (37.9)	462 (31.6)
70 or older	92 (25.6)	209 (14.3)
Missing values	0 (0.0)	2 (0.1)
Province		
Newfoundland	11 (3.0)	117 (8.0)
Prince Edward Island	5 (1.4)	138 (9.4)
Nova Scotia	64 (17.8)	236 (16.1)
Manitoba	60 (16.7)	153 (10.5)
Saskatchewan	58 (16.2)	124 (8.5)
Alberta	86 (24.0)	291 (19.9)
British Columbia	75 (20.9)	402 (27.5)
Missing values	0 (0.0)	4 (0.3)
Education (in years)		
≤ 9	79 (22.2)	266 (18.3)
10 – 12	170 (47.7)	664 (45.8)
>12	107 (30.1)	520 (35.9)
Missing values	3 (0.8)	15 (1.0)
Marital status		
Common law	8 (2.2)	51 (3.5)
Divorced/Separated	31 (8.6)	140 (9.6)
Married	223 (62.1)	960 (65.7)
Single	29 (8.0)	70 (4.8)
Widowed	68 (18.9)	236 (16.1)
Other	-	4 (0.3)
Missing values	0 (0.0)	4 (0.3)
Occupational group		
Non-hazardous	288 (80.7)	1,275 (87.3)
Hazardous	69 (19.3)	185 (12.7)
Missing values	2 (0.5)	5 (0.3)

[§] Due to rounding errors, percentages (denominators exclude missing values) of the observed values may not add up to 100

Table 15 Frequencies of female cases and controls with respect to the categories of cigarette smoking, alcohol, and coffee consumption

Characteristic	Cases (n = 359) Number (%)[§]	Controls (n = 1,465) Number (%)
Smoked 100+ cigarettes		
Yes	260 (72.4)	753 (51.4)
No	99 (27.6)	712 (48.6)
Smoking status[*]		
Never smoked	99 (27.6)	712 (48.6)
Ex-smoker	110 (30.7)	461 (31.5)
Current smoker	150 (41.8)	292 (19.9)
Pack-years of smoking[**]		
0	99 (28.0)	712 (49.3)
1 – 4	26 (7.4)	187 (13.0)
5 – 11	36 (10.2)	177 (12.3)
12 – 23	83 (23.5)	190 (13.2)
≥24	109 (30.9)	177 (12.3)
Missing values	6 (1.7)	22 (1.5)
Coffee consumption (per day)		
Never or less than 1 cup	91 (25.7)	480 (33.4)
1 cup or more	263 (74.3)	958 (66.6)
Missing values	5 (1.4)	27 (1.8)
Alcohol consumption (servings per week)[***]		
0	195 (54.3)	709 (48.5)
>0 to ≤ 0.9	33 (9.2)	163 (11.1)
>0.9 to ≤ 2.0	38 (10.6)	235 (16.1)
>2.0 to ≤6.0	34 (9.5)	184 (12.6)
>6.0	59 (16.4)	172 (11.8)
Missing values	0 (0.0)	2 (0.1)

[§] Due to rounding errors, percentages (denominators exclude missing values) of the observed values may not add up to 100

* Two years before the interview

** Quartiles based on the distribution in smoker male controls

*** Combined consumption: beer, wine, and liquor. Quartiles based on the distribution of alcohol consumption in male drinker controls

51.0% of the controls had ever smoked 100 or more cigarettes. The prevalence of current smokers amongst cases was approximately twice that amongst controls (41.8% vs. 20.0%). The proportions of women who had 24 or more pack-years of cigarette smoking in cases and controls were 30.9% and 12.3%, respectively. Coffee consumption (1 cup or more per day) was more prevalent amongst cases (74.3% vs. 66.6%). In cases, the proportion of women who had consumed alcohol of 6.0 servings per week or more was higher than that in controls (16.4 vs. 11.8 servings per week, respectively).

Tables 16-17 display distributions of female cases and controls with respect to their reproductive factors. According to Table 16, cases were more likely to be nulliparous and postmenopausal. Cases and controls were similar in their breast-feeding history. Controls compared to cases were more likely to have greater number of live births.

Distributions of age at first and last menstruation, years menstruated, and age at first pregnancy, are displayed in Table 17. The means and standard deviations of these distributions amongst cases and controls were similar. Note that since years menstruated was calculated using different formula for pre- and post-menopausal women, distribution of this variable is presented separately for pre- and post-menopausal women.

Table 16 Frequencies of selected reproductive factors for female cases and controls

Characteristic	Cases (n = 359) Number (%)[§]	Controls (n = 1,465) Number (%)
Ever pregnant		
Yes	305 (85.2)	1,298 (88.8)
No	53 (14.8)	163 (11.2)
Missing values	1 (0.3)	4 (0.3)
Parity		
Nulliparous	65 (18.2)	201 (13.8)
Parous	293 (81.8)	1,255 (86.2)
Missing values	1 (0.3)	9 (0.6)
Number of live births		
0	65 (18.2)	201 (13.8)
1 - 2	118 (33.0)	523 (35.9)
3 – 4	114 (31.8)	499 (34.3)
≥5	61 (17.0)	233 (16.0)
Missing values	1 (0.3)	9 (0.6)
Menopausal status		
Post-menopausal	306 (85.3)	1,003 (68.7)
Pre-menopausal	53 (14.7)	458 (31.3)
Missing values	0 (0.0)	4 (0.3)
Still menstruate		
Yes	31 (8.6)	380 (26.0)
No	328 (91.4)	1,081 (74.0)
Missing values	0 (0.0)	4 (0.3)
Breastfeeding		
Ever	235 (65.8)	940 (64.3)
Never	122 (34.2)	521 (35.7)
Missing values	2 (0.6)	4 (0.3)

[§] Due to rounding errors, percentages (denominators exclude missing values) of the observed values may not add up to 100

Table 17 The distribution of age and selected reproductive factors amongst female cases and controls

Characteristic*	Cases (n = 359)			Controls (n = 1,465)		
	Range	Mean (SD)**	Missing #	Range	Mean (SD)	Missing
Age at interview						
Pre-menopausal sample	22 – 58	43.5 (9.0)	0 (0.0)	20 – 58	42.7 (8.5)	2 (0.4)
Post-menopausal sample	40 – 75	64.2 (7.8)	0 (0.0)	27 – 76	62.3 (7.9)	0 (0.0)
Age at first menstruation	8 – 18	13.0 (1.6)	27 (7.5)	8 – 19	12.9 (1.6)	118 (8.0)
Age at last menstruation†	19 – 58	44.8 (7.4)	20 (6.3)	17 – 58	45.7 (7.0)	63 (6.0)
Years menstruated ‡						
Combined sample §	3 – 44	31.1 (7.9)	46 (12.8)	2 – 47	31.7 (7.8)	165 (11.3)
Pre-menopausal sample	6 – 44	30.1 (9.4)	5 (9.3)	6 – 47	29.9 (8.6)	38 (8.3)
Post-menopausal sample	3 – 44	31.3 (7.6)	41 (13.4)	2 – 47	32.5 (7.2)	127 (12.6)
Age at first pregnancy ¶	14 – 41	23.6 (4.5)	18 (6.1)	13 – 44	23.5 (4.6)	32 (2.5)

* The unit of measurement for all variables in this table is 'in years'

** Standard deviation

\# Number and percentage (%) of missing values (4 controls with unknown menopausal status were excluded from the calculation of the percentage of missing values)

† Sample of postmenopausal women (306 cases and 1,003 controls)

† ‡Percentages of missing values for pre- and post-menopausal women are based on 53 cases/458 controls and 306 cases/1,003 controls, respectively

§ Combined post- and pre-menopausal sample of women

¶ Duration of pregnancy ≥ 5 months (percentages are based on the sample of parous women: 293 cases and 1,255 controls)

6.3.2. Total Physical Activity Score and Body Mass Index across the Subsets of Female Controls

The distribution of total physical activity scores across different subsets of the female controls is presented in Table 18 and the corresponding box-plots (see Appendix B2). According to the observed patterns, more physically active women tended to be younger, more educated, and to have a higher level of income. There was a decreasing trend in the total physical activity score moving from never to heavy smokers (≥ 24 pack-years).

Table 18 Summary statistics for total physical activity score across age, smoking, educational, and income adequacy groups in female controls, NECSS study, Canada, 1994-97

Group	Mean	25%	Median	75%	N *
Age (years)					
< 50	83.3	24.2	60.3	126.3	416
50 – 70	81.2	23.5	62.5	118.7	826
> 70	64.6	7.1	43.7	94.1	164
Smoking (pack-years)					
0	79.7	23.5	58.7	118.5	685
1 – 4	85.3	24.1	70.1	125.1	184
5 – 11	84.7	26.2	72.2	129.7	171
12 – 23	83.4	27.3	57.6	121.0	181
≥ 24	69.5	18.4	47.9	101.9	167
Education (years)					
0 – 9	59.3	6.9	31.2	91.0	255
10 – 12	78.8	21.9	57.6	120.5	640
≥ 13	92.8	35.0	77.9	134.6	497
Income adequacy					
Low	66.8	11.7	43.7	102.9	283
Lower middle	75.9	20.2	60.1	114.2	262
Upper middle	83.4	33.7	64.9	123.2	343
High	92.3	37.0	77.1	137.7	166

* Sample size

According to the distribution of body mass index amongst the subsets of female controls, displayed in Table 19 and the corresponding box-plots (see Appendix B2), women with higher body mass index were more likely to diet during their adult lives, to have greater number of live births, lower income, and lower education. In summary, the observed distributional patterns of the main exposure variables amongst the subsets of female controls conformed to those expected in the general population.

Table 19 Summary statistics for body mass index across parity, diet, income adequacy, and educational groups in female controls, NECSS study, Canada, 1994-97

Group	Mean	25%	Median	75%	N*
Number of live births					
0	25.0	21.3	23.5	27.5	198
1 – 2	24.8	21.7	23.8	27.1	520
3 – 4	25.2	22.3	24.4	27.3	497
≥ 5	26.3	23.1	25.8	29.2	230
Number of times dieted ‡					
Never	22.9	20.6	22.6	24.7	504
1 – 2	24.9	22.1	24.1	27.0	337
3 – 5	26.4	23.2	25.9	28.4	267
6 – 8	27.3	23.5	26.4	29.9	136
9 – 11	28.3	24.5	27.3	31.2	36
≥12	29.1	24.7	28.2	32.9	152
Income adequacy					
Low	26.4	22.5	24.7	29.3	294
Lower middle	25.3	21.9	24.2	27.4	271
Upper middle	24.7	21.9	24.0	26.7	355
High	24.6	21.9	23.5	26.4	166
Education (years)					
0 – 9	26.1	22.6	25.5	28.4	259
10 – 12	25.6	22.2	24.6	27.9	663
≥ 13	24.4	21.3	23.5	26.4	517

* Sample size

‡ The number of times dieted during adult life

6.3.3. Model-Based Analysis

The age- and Province-adjusted ORs for recreational physical activity and body mass index levels in relation to bladder cancer are shown in Table 20. The ORs for the upper two quartiles (> 60.3 to ≤124.0 and >124.0 MET-hours/month) of total physical activity score were decreased relative to those of the reference category. However, these estimates reached only marginal statistical significance (OR = 0.70, 95% CI: 0.50, 1.00 and OR = 0.74, 95% CI: 0.52, 1.05). The test of trend for physical activity was statistically significant at $\alpha = 0.05$ level (P value = 0.03). The ORs quantifying the strength of the association between body mass index and bladder cancer risk were close to unity for all the categories of body mass index. The corresponding 95% CIs and P value for trend (P value = 0.25) were not suggestive of any statistically significant differences in the risk between the reference and upper categories of body mass index. Women who ever had been employed in any hazardous occupation(s) for at least one year had an elevated bladder cancer risk relative to those who had never had such occupation(s) (OR = 1.56, 95% CI: 1.14, 2.15). The age- and Province-adjusted OR for education was not statistically significant. The test for trend for education yielded a large P value suggesting the absence of trend in risk gradient (see Table 20).

Almost all age- and Province-adjusted ORs associated with smoking were large in magnitude and highly statistically significant (see Table 21). For example, current smokers had approximately 4.5 times the risk of that for never smokers. The age-and Province-adjusted risk for bladder cancer amongst smokers in the highest category of pack-years relative to never smokers was 4.75 (95% CI: 3.39, 6.66).

Table 20 Age* - and Province-adjusted odds ratios for recreational physical activity, body mass index, and other characteristics in female cases (n = 359) and controls (n = 1,465), NECSS study, Canada, 1994-97

Characteristic	OR	95% CI †	Number of subjects** Cases/Controls
Total physical activity score (MET-hours/month) ‡			344/1,413
≥0 to ≤ 22.5	1.00¶		
> 22.5 to ≤60.3	0.89	0.64, 1.25	
>60.3 to ≤124.0	0.70	0.50, 1.00	
>124.0	0.74	0.52, 1.05	
P value for trend	0.03		
Body mass index (kg/m^2) §			358/1,448
< 22.0	1.00¶		
≥22.0 to < 24.3	0.98	0.68, 1.40	
≥24.3 to < 27.4	1.04	0.73, 1.47	
≥27.4	1.18	0.84, 1.67	
P value for trend	0.25		
Occupational group			357/1,454
Non-hazardous	1.00¶		
Hazardous	1.56	1.14, 2.15	
Education (in years)			356/1,444
≤ 9	1.00¶		
10 – 12	0.93	0.67, 1.28	
>12	0.85	0.60, 1.22	
P value for trend	0.39		

* Age categorized as 20-39, 40-49, 50-59, 60-69, and 70 years or older

** Subjects with missing values for the variable were excluded from the analysis

† 95 percent confidence interval

‡ Quartiles based on the distribution of total physical activity score in female controls

§ Quartiles based on the distribution of body mass index in female controls

¶ Reference category

The result based on the test of trend for the relationship between pack-years of cigarette smoking and risk of bladder cancer was highly statistically significant (P value < 0.01).

Table 21 Age* - and Province-adjusted odds ratios for smoking, coffee and alcohol consumption groups in female cases (n = 359) and controls (n = 1,465), NECSS study, Canada, 1994-97

Characteristic	OR	95% CI †	Number of subjects** Cases/Controls
Smoked 100+ cigarettes			359/1,459
No	1.00¶		
Yes	2.69	2.06, 3.50	
*Smoking status****			359/1,459
Never smoked	1.00¶		
Ex-smoker	1.74	1.28, 2.37	
Current smoker	4.46	3.28, 6.07	
Pack-years of smoking ‡			353/1,437
0	1.00¶		
1 – 4	1.13	0.70, 1.82	
5 – 11	1.60	1.04, 2.46	
12 - 23	3.38	2.39, 4.79	
≥24	4.75	3.39, 6.66	
P value for trend	< 0.01		
Coffee consumption (per day)			354/1,434
Never or less than 1 cup	1.00¶		
1 cup or more	1.14	0.85, 1.51	
Alcohol consumption (servings per week) §			359/1,459
0	1.00¶		
>0 to ≤0.9	0.74	0.48, 1.12	
>0.9 to ≤2.0	0.61	0.41, 0.88	
>2.0 to ≤6.0	0.70	0.46, 1.06	
>6.0	1.22	0.86, 1.73	
P value for trend	0.18		

* Age categorized as 20-39, 40-49, 50-59, 60-69, and 70 years or older

** Subjects with missing values for the variable were excluded from the analysis

*** Two years before the interview

† 95 percent confidence interval

‡ Quartiles based on the distribution of pack-years in male smoker controls

§ Combined consumption: beer, wine, and liquor. The quartiles were based on the distribution of alcohol consumption in male drinker controls

¶ Reference category

Table 21 also displays age- and Province-adjusted ORs and 95% CIs for coffee and alcohol consumption. Based on the age- and Province-adjusted OR of 1.14 and 95% CI: 0.85, 1.51, coffee consumption was not associated with risk of bladder cancer. The age- and Province-adjusted ORs for alcohol consumption and education were not statistically significant. The test for trend for alcohol consumption yielded a large P value suggesting the absence of trend in the risk gradient.

The age- and Province-adjusted measures of association for the reproductive factors are shown in Table 22. Of these variables, parity (being parous or having at least one live birth vs. being nulliparous) and the number of years menstruated (only for post-menopausal women) were statistically significantly associated with risk of bladder cancer. For example, parous women relative to nulliparous women had 40% reduced risk of bladder cancer (OR = 0.59, 95% CI: 0.42, 0.81). In post-menopausal women, the age- and Province-adjusted OR for having menstruated 33 or more years, as compared to those who had menstruated less than 33 years was 0.67 (95% CI: 0.50, 0.89).

The final multivariable model consisting of total physical activity score, body mass index, occupation, the number of live births, the number of pack-years of cigarette smoking, age, and Province is shown in Table 23. The OR-s for the two upper quartiles (> 60.3 to ≤ 124.0 and >124.0 MET-hours/month) of total physical activity score were not statistically and significantly different from that of the reference quartile (OR = 0.70, 95% CI: 0.48, 1.02 and OR = 0.76, 95% CI: 0.52, 1.21).

Table 22 Age* - and Province-adjusted odds ratios for reproductive factors in female cases (n = 359) and controls (n = 1,465), NECSS study, Canada, 1994-97

Characteristic***	OR	95% CI †	Number of subjects** Cases/Controls
Ever pregnant			358/1,459
No	1.00¶		
Yes	0.62	0.43, 0.88	
Parity			358/1,454
Nulliparous	1.00¶		
Parous	0.59	0.42, 0.81	
Breastfeeding			357/1,459
Never	1.00¶		
Ever	1.00	0.77, 1.30	
Number of live births			358/1,454
0	1.00¶		
1 – 2	0.63	0.43, 0.90	
3 – 4	0.54	0.37, 0.78	
≥5	0.60	0.39, 0.92	
P value for trend		0.03	
Age at first menstruation			332/1,345
≤ 13	1.00¶		
> 13	0.98	0.75, 1.28	
Age at last menstruation §			286/939
≤ 45	1.00¶		
> 45	0.72	0.54, 0.95	

Continued...

Table 22 Age* - and Province-adjusted odds ratios for reproductive factors in female cases (n = 359) and controls (n = 1,465), NECSS study, Canada, 1994-97 (continued)

Characteristic***	OR	95% CI †	Number of subjects** Cases/Controls
Years menstruated #			
Pre-menopausal sample			48/419
< 33	1.00¶		
≥33	0.65	0.22, 1.89	
Post-menopausal sample			265/878
< 33	1.00¶		
≥33	0.67	0.50, 0.89	
Age at first pregnancy ‡			293/1,253
≤ 23	1.00¶		
> 23	1.05	0.80, 1.38	

* Age categorized as 20-39, 40-49, 50-59, 60-69, and 70 years or older
** Subjects with missing values for the variable were excluded from the analysis
*** Variables in this table are dichotomized (except for 'years menstruated') at the median of their distribution in controls
† 95 percent confidence interval
§ Sample of postmenopausal women
# Dichotomization based on log odds curve (smooth log odds versus 'years menstruated')
‡ Sample of parous women
¶ Reference category

Test for trend for this association was not statistically significant (P value = 0.14), suggesting the absence of linear gradient in risk across the quartiles of physical activity.

Body mass index was not associated with risk of bladder cancer. The adjusted odds of the outcome for women in the highest body mass index quartile (≥ 27.4) were not statistically significantly different from those for women in the lowest quartile of body mass index (reference category) (OR = 1.07, 95% CI: 0.74, 1.55).

Table 23 Multivariable odds ratios for female bladder cancer based on 334 cases and 1,377 controls, NECSS, Canada, 1994-97

Characteristic	OR[#]	95% CI [†]
Total physical activity score (MET-hours/month) [‡]		
≥0 to ≤ 22.5	1.00[¶]	
> 22.5 to ≤ 60.3	0.88	0.61, 1.26
> 60.3 to ≤ 124.0	0.70	0.48, 1.02
> 124.0	0.76	0.52, 1.21
P value for trend	0.14	
Body mass index (kg/m^2)[*]		
< 22.0	1.00[¶]	
≥ 22.0 to < 24.3	1.09	0.75, 1.60
≥ 24.3 to < 27.4	1.03	0.71, 1.50
≥ 27.4	1.07	0.74, 1.55
P value for trend	0.81	
Number of live births		
0	1.00[¶]	
1 – 2	0.59	0.40, 0.88
3 – 4	0.52	0.35, 0.77
≥5	0.55	0.35, 0.89
P value for trend	0.03	
Occupational Group		
Non-hazardous	1.00[¶]	
Hazardous	1.55	1.10, 2.19
Pack-years of smoking [§]		
0	1.00[¶]	
1 – 4	1.13	0.69, 1.86
5 – 11	1.63	1.05, 2.55
12 – 23	3.41	2.37, 4.90
≥ 24	4.78	3.37, 6.78
P value for trend	< 0.01	

[#] Odds ratio adjusted for all the variables listed in this table (except for the variable under consideration), age (20-39, 40-49, 50-59, 60-69, and 70 years or older), and Province

[†] 95 percent confidence interval

[‡] Quartiles based on the distribution of total physical activity score in female controls

[*] Quartiles based on the distribution of body mass index in female controls

[§] Quartiles based on the distribution of pack-years in female smoker controls

[¶] Reference category

Test for trend did not suggest any trends in magnitudes of the ORs across the body mass index quartiles (P value for trend = 0.81).

Having at least 1 live birth was statistically significantly associated with a decreased bladder cancer risk. For example, the odds of having bladder cancer for women with one or two live births, relative to nulliparous women, were reduced by approximately 40% (OR = 0.59, 95% CI: 0.40, 0.88). The additional adjustment for confounding did not change the magnitude of the coefficient for occupation (OR = 1.56 in Table 20 vs. OR = 1.55 in Table 23). Moreover, the estimate maintained its statistical significance in the final multivariable model (95% CI: 1.10, 2.19). The strong and positive association between the number of pack-years of cigarette smoking and bladder cancer risk persisted in the multivariable model (P value for trend < 0.01).

The stratification by menopausal status in order to explore potential effect modification was not completely successful. Because of the small sample size, the logistic model based on the pre-menopausal women could not be investigated (very unstable estimates of doubtful validity). The model had such numerical problems as low number of events per variable (EPV < 10) and zero cells in the quartiles of total physical activity score and body mass index [87; 93]. The multivariate logistic regression model for the sample of post-menopausal women is presented in Table 24.

Table 24 Multivariable odds ratios for female bladder cancer based on the post-menopausal sample of 284 cases and 938 controls, NECSS, Canada, 1994-97

Characteristic	OR[#]	95% CI [†]
Total physical activity score (MET-hours/month) [‡]		
≥0 to ≤ 22.5	1.00[¶]	
> 22.5 to ≤ 60.3	0.87	0.58, 1.30
> 60.3 to ≤ 124.0	0.77	0.51, 1.16
> 124.0	0.72	0.47, 1.11
P value for trend	0.13	
Body mass index (kg/m^2)[*]		
< 22.0	1.00[¶]	
≥ 22.0 to < 24.3	0.92	0.60, 1.41
≥ 24.3 to < 27.4	0.85	0.56, 1.29
≥ 27.4	0.79	0.52, 1.19
P value for trend	0.24	
Number of live births		
0	1.00[¶]	
1 – 2	0.53	0.34, 0.85
3 – 4	0.48	0.31, 0.75
≥5	0.50	0.30, 0.83
P value for trend	0.03	
Occupational Group		
Non-hazardous	1.00[¶]	
Hazardous	1.60	1.10, 2.34
Pack-years of smoking[§]		
0	1.00[¶]	
1 – 4	0.80	0.42, 1.48
5 – 11	1.70	1.03, 2.80
12 – 23	3.69	2.45, 5.55
≥ 24	5.00	3.46, 7.35
P value for trend	< 0.01	

[#] Odds ratio adjusted for all the variables listed in this table (except for the variable under consideration), age (20-39, 40-49, 50-59, 60-69, and 70 years or older), and Province

[†] 95 percent confidence interval

[‡] Quartiles based on the distribution of total physical activity score in female controls

[*] Quartiles based on the distribution of body mass index in female controls

[§] Quartiles based on the distribution of pack-years in female smoker controls

[¶] Reference category

In general, model-based results for the total (combined pre- and post-menopausal women) and post-menopausal samples of women were similar.

In the sub-group of parous women, the multivariable logistic regression estimates of ORs for recreational physical activity, body mass index, and occupation were very similar to those already observed and presented in Table 23 (data not shown). The number of live births was not associated with risk of bladder cancer. For example, after adjusting for recreational physical activity, body mass index, occupation, cigarette smoking, age, and Province, the odds of having bladder cancer in women with 3 - 4 live births and ≥ 5 live births were similar to those in women with 1-2 live births - the reference category (OR = 0.87, 95% CI: 0.63, 1.20 and OR = 0.92, 95% CI: 0.61, 1.37, respectively). This finding suggested that the earlier observed association for the number of live births in relation to bladder cancer risk, presented in Tables 23 - 24, may have reflected simply the effect of being parous relative to nulliparous. Indeed, the relative odds of having bladder cancer for the three parity groups (1 - 2, 3 - 4, and ≥ 5) in Table 23 were similar in magnitude: 0.59, 0.52, and 0.55, respectively. The likelihood ratio test result also confirmed that the number of live births did not contribute to the risk of bladder cancer. Specifically, the deviances of the models containing the number of live births (0, 1 - 2, 3 - 4, and ≥ 5) and parity (parous vs. nulliparous) were very similar, 1,449 and 1,450, respectively. Given the above-mentioned, it was decided to replace the number of live births (0, 1 - 2, 3 - 4, and ≥ 5) with the dichotomous variable defining parous (1 - 2, 3 - 4, and ≥ 5 live births) vs. nulliparous women. The multivariable odds ratio for being parous vs. nulliparous was 0.56 (95% CI: 0.39, 0.80). This replacement did not affect the previously observed

magnitudes of ORs and 95% CIs for rest of the variables (Tables 23-24). The Hosmer-Lemeshow goodness-of-fit test for this final model yielded a Chi-square (8 degrees of freedom) of 7.60 (P value = 0.47), indicating a satisfactory fit of the model.

In the sub-group of post-menopausal women, the corresponding estimate of the odds ratio (parous vs. nulliparous) was OR = 0.50 (95% CI: 0.33, 0.76). The Hosmer-Lemeshow goodness-of-fit test for the final model yielded a Chi-square (8 degrees of freedom) of 9.20 (P value = 0.32).

7. Discussion

7. 1. The Sample of Men

7. 1. 1. Main Findings

In this study there was no relationship between the levels of recreational physical activity and risk of bladder cancer in men. None of the estimates of OR for recreational physical activity was statistically significant (see Table 12). Note that the estimate of OR for the last upper quartile of physical activity (> 150.0 MET- hours/month) was slightly elevated relative to the reference category. But the overall pattern of data was not indicative of a linear dose-response relationship. The test for trend yielded the P value of 0.13, suggesting the absence of a linear gradient in risk associated with levels of recreational physical activity.

The approximate amount of activity corresponding to the quartiles of total physical activity score could be interpreted by using median of each quartile. For example, based on the median score of total physical activity, an average annual amount of recreational physical activity for a subject in the reference category (the lowest quartile) of total physical activity (≥ 0 to ≤ 24.2 MET-hours/month) would be to perform any single activity with MET ≤ 5 (e.g., walking, bowling, home exercise, or social dancing), 3 – 6 times per week, for about 15-30 minutes each session, in one season (i.e., winter, spring, summer, or fall). An average annual amount of recreational physical activity for a subject in the second upper quartile (> 24.2 to ≤ 73.6 MET- hours/month) would be to perform any single activity with MET ≤ 5 (e.g., walking, bowling, home exercise, or social dancing), 3 – 6 times per week, for about 31 - 60 minutes each session, in two seasons.

For a subject to fall in the third upper quartile (>73.6 to ≤ 150 MET-hours/month) of total physical activity score, he would have to perform any single activity with MET > 5 (e.g., jogging, tennis, skiing or swimming), 3 - 6 times per week, for at least 60 minutes each session, in three seasons. An average annual amount of the physical activity for the highest quartile (> 150.0 MET-hours/month) of total physical activity score would be to perform any single activity with MET > 5 (e.g., jogging, tennis, skiing or swimming), every day, for at least 60 minutes each session, in three or four seasons (see Table 2).

The results of this study indicated that relative body weight was not associated with risk of bladder cancer in men. Although the odds of bladder cancer in the final model were elevated for the third and fourth quartiles of body mass index, only one of the corresponding ORs was statistically significant. Besides, there was no clear trend of risk gradient across the quartiles of BMI (P value for trend was 0.10). The magnitude of OR for the middle quartile (≥ 25.7 to < 28.0 kg/m^2) was larger (OR = 1.49) than that (OR = 1.29) for the upper quartile of BMI (≥28.0 kg/m^2), rendering this observed pattern biologically less likely scenario. This association is more likely to reflect pattern of non-differential misclassification of BMI. The overall evidence was not strong enough to reject the Null Hypothesis of 'no association' and to conclude that men with higher BMI were at an increased risk of bladder cancer.

A marginal but statistically significant association was found between occupation and risk of bladder cancer. Men who had ever held any 'high risk' occupation(s) were at increased risk of having bladder cancer compared to those who had never held such

occupation(s) (OR = 1.22, 95% CI: 1.00, 1.50). Because the occupational classification was a simple dichotomy, the observed OR of 1.22 could be an underestimate of the true odds ratio due to some amount of non-differential misclassification. The effect of coffee consumption was marginally statistically significant, possibly indicating a 30% increase in risk for bladder cancer amongst drinkers of 1 cup or more a day relative to never drinkers. Cigarette smoking was strongly associated with an increased risk of bladder cancer.

Of the variables considered in the 'bivariable' (age- and Province-adjusted) analyses, education and alcohol consumption were not kept in the final model for several reasons. The estimates of multivariable (adjusted for the main exposure variables, occupation, coffee consumption, pack-years of smoking, age, and Province) ORs for the educational categories 10 – 12 and > 12 years versus ≤ 9 years were 1.00 (95% CI: 0.77, 1.29) and 0.87 (0.66, 1.15), respectively. The likelihood ratio test yielded a P value greater than 0.10, suggesting that education did not contribute significantly towards the model's overall fit. The inclusion/exclusion of education in the model did not influence either the magnitudes or confidence intervals of other regression coefficients. Similarly, there were no associations between the categories of alcohol consumption and bladder cancer risk. The multivariable ORs for the quartiles of alcohol consumption were close to unity and none of them was statistically significant at $\alpha = 0.05$. Likelihood ratio test for this variable yielded a P value greater than 0.10. The exclusion of alcohol consumption from the model did not influence the magnitudes of the regression coefficients, suggesting that it was not a confounder.

7. 1. 2. Consistency of the Findings with Those of Other Studies

Our finding of no association between physical activity levels and bladder cancer risk in men are in agreement with four out of the five previously reported studies [35-37; 40; 79]. Only one study (prospective cohort design) found that vigorous recreational activity was associated with increased risk of bladder cancer (RR = 2.06; 95% CI: 1.08, 3.95) [37]. The different results in the two studies with respect to the effects of physical activity on bladder cancer cannot be readily explained. Different study designs, methods of measurement of physical activity, sample size, and adjustment strategies might have contributed to this discrepancy. Since Wannamethee et al. [37] obtained their physical activity data before the occurrence of bladder cancer, these measurements would not have been influenced by the presence of disease. However, because these measurements were taken only at baseline, some changes in patterns of activity might have taken place during the follow-up. The participation/response rate in this prospective study was high (90%), rendering the influence of selection bias minimal. The RR was adjusted for cigarette smoking, BMI, socioeconomic status (SES), and alcohol intake (not for occupation). However, it was based on a small sample of 92 bladder cancer cases. The authors stated that some cases (the actual proportion not reported) were ascertained by mailed questionnaires. If the proportion of cases ascertained by mailed questionnaires was substantial, the possibility of the differential rates of cancer ascertainment across the physical activity levels should not be excluded. This could happen if physically more active participants were of higher SES and were more likely to report bladder cancer than those of lower SES. The over-ascertainment of the disease amongst physically more

active subjects could inflate the relative risk, suggesting a positive association between physical activity and bladder cancer risk [87].

Another prospective cohort study with a larger sample size reported that levels of physical activity were not related to risk of bladder cancer (heavy vs. sedentary activity RR = 1.05; 95% CI: 0.60, 1.84) [36]. Although this study reported that the estimate was controlled for smoking, it did not describe how this variable was measured and categorized. The authors did not control the analysis for the effect of occupation.

The association between BMI and bladder cancer risk found in this study are consistent with those from the four studies [39; 41; 77-78]. One study demonstrated a statistically elevated risk for the relationship between body mass index and bladder cancer risk (SIR = 1.3, 95% CI: 1.0, 1.6)[78]. However, the magnitude of SIR is small and its statistical significance is marginal. The adjustment of this estimate was restricted only to age and gender.

The results of this thesis regarding the harmful effects of cigarette smoking and certain 'high risk' occupation(s) in relation to bladder cancer risk were consistent with those from the previous research [2-3; 5; 7; 44]. Slightly increased risk for the coffee consumption, found in this study, is in agreement with previous studies in men [55].

7. 2. The Sample of Women

7. 2. 1. Main Findings

The ORs for the quartiles of total physical activity score in the final model were not statistically significantly different from that for the reference quartile (Table 23). The OR for the middle quartile was 0.70, towards the direction of protective effect, but its magnitude was more extreme than that for the upper quartile (OR = 0.76). The high P value (P = 0.14) for trend did not support a linear decrease in risk. Overall, the evidence was not convincing enough to conclude that the levels of recreational physical activity were inversely associated with risk of female bladder cancer.

There was no association between body mass index and female bladder cancer risk.

This study found that parous women (having at least 1 live birth) had about a 40 - 45% reduction in risk of bladder cancer compared to nulliparous women. For parous women, having a greater number of live births did not confer any additional protection against bladder cancer.

Although statistically marginally significant, the estimate of OR for occupation suggested a 55% elevation in risk of bladder cancer for women who had ever held 'high-risk' occupation(s) compared to those who never held such occupation(s). Similarly to men, cigarette smoking in women was strongly associated with an increased risk of bladder cancer. In the final model, the risk gradient for the association between cigarette smoking

and bladder cancer risk conformed to dose-response relationship (P value for trend < 0.01).

Of the variables considered in the 'bivariable' (age- and Province-adjusted) analyses, education, coffee consumption, alcohol consumption, breastfeeding, and the number of years menstruated were not kept in the final multivariable model (see Table 23). The addition of these variables to the model, one at a time, did not contribute significantly towards the model's overall fit (likelihood ratio test-based P value > 0.10). Moreover, their inclusion/exclusion from the model did not influence the magnitudes and/or confidence intervals of other regression coefficients.

The logistic regression models based on the total (pre- and post-menopausal) and post-menopausal samples of women were similar (Tables 23-24). This may have reflected the fact that post-menopausal women constituted most of the total sample of women (306/359 cases and 1,003/1,465 controls).

7. 2. 2. Consistency of the Findings with Those of Other Studies

The extensive literature search identified and retrieved only one study exploring the relationship between leisure-time physical activity and bladder cancer risk in women [38]. This was a 13-year follow-up study of 37,459 post-menopausal women. The study authors used a physical activity index classifying the study subjects into 'low', 'moderate', and 'high' activity categories. The age-adjusted RRs and 95% CIs for bladder cancer risk among 'moderate' and 'high' activity groups relative to 'low' activity group

were 0.62 (95% CI: 0.38, 1.00) and 0.73 (95% CI: 0.46, 1.17), respectively. The study authors were unable to draw definitive conclusions regarding the protective effect of physical activity because of the absence of trend in risk and the statistically non-significant RRs. The patterns of risk (the direction of effect, the magnitudes of RR, and 95% CIs) observed in this study are similar to those obtained in the present study (see Table 23).

Four out of the five studies reporting the associations between body mass index and bladder cancer risk found no association between the two factors [22; 38; 41; 77-78]. The results of this thesis are consistent with those of the four studies [22; 38; 77-78]. A Swedish study, using directly standardized rates, found that women having BMI greater than 28.6 kg/m^2 were at a 40% increased risk of having bladder cancer compared to the Swedish general population of women of the same age (SIR = 1.40, 95% CI: 1.0, 1.9) [41].

In the retrieved literature, only three studies reported on the relationship between parity and bladder cancer [10; 22; 38]. Only one of these studies found an association (parous vs. nulliparous OR = 0.67, 95% CI: 0.44, 1.00) [10]. Another study reported an estimate of relative risk suggesting a protective effect of multi-parity (30% reduction in risk for those with 3 or more live births). However, the RR was not statistically significant (RR = 0.70; 95% CI: 0.38, 1.27). This was probably due to a small number of cancer events and a low study power (n = 100) [38]. Only one study reported the relationships between female reproductive factors and bladder cancer risk in parous women [10]. The results of that study

and this thesis agreed that there was no association between the number of live births and bladder cancer risk in parous women.

The findings of this thesis regarding the harmful effects of cigarette smoking and certain 'high risk' occupation(s) in relation to risk of female bladder cancer were consistent with those from previous research [2-3; 5; 7; 44]

In summary, the associations found in this study of women are in agreement with the results from majority of the previously conducted studies on the topic.

7. 3. The Role of Bias

This section will review different types of bias that may affect results of case-control studies. The possibility for the presence and extent of the influence of each type of bias on the study results is considered.

Non-Differential Misclassification

Non-differential misclassification in reporting physical activity and body mass index would have been introduced due to the study subjects' inaccurate recall of their activity patterns and body weight/height (difficulties in recall of past events and/or subjectivity of perception) and/or low reproducibility of the physical activity questionnaire. High correlation shown between self-reported and measured body weight/height ($r = 0.95$) [28] would not prevent the occurrence of non-differential misclassification with respect to body mass index (a measure could be valid but at the same time imprecise).

In the majority cases of dichotomous exposure measurement, the effect of random error would be to decrease the ability of the study to detect an effect of exposure on risk of the outcome by attenuating an estimate of a true OR [94]. However, when a polychotomous exposure is used, as in case of this study (i.e., quartiles of physical activity and body mass index), the direction of bias is less predictable, i.e., the OR estimates may be biased towards or away from the null [95-97]. For example, the observed magnitudes of odds ratio for BMI in men and physical activity in women could have been inflated partially or entirely due to non-differential misclassification.

Since occupation was defined rather crudely as the dichotomy, some amount of non-differential misclassification could have occurred. Namely, some truly exposed subjects may have been misclassified as non-exposed and vice versa. Certain subjects would have been misclassified with respect to the occupational exposure due to the lack of more detailed information regarding their exact job-related duties. For example, someone who worked at a rubber factory as a manager could have been classified as exposed. The relevant literature [5] lists many other 'hazardous' occupations (i.e., mechanics, nursery workers, photographic workers, security guards, sailors, or bootblacks) for bladder cancer which were not considered as 'exposure' in this study. Therefore, some subjects who had reported to have held these occupations might have been classified as 'non-exposed'. But the findings for these occupations are less convincing and besides there were very few subjects in the study sample who held these types of occupation. Another source of misclassification would have been introduced by misclassifying the truly non-exposed study subjects as exposed if the time period of their 'hazardous' occupation was outside

the etiologically relevant time-window (i.e., after a life-time managerial work, he/she became a farm worker 1-5 years prior the study interview). The effect of non-differential misclassification of a dichotomous exposure is to bias the estimated OR towards the null. Thus, in both men and women, the observed ORs for occupation may have underestimated the true effect size related to the employment in 'high risk' occupation(s).

Differential Misclassification

Case-control studies are prone to recall bias - a common type of differential misclassification. Differential misclassification can either overestimate (inflate) or underestimate (dilute/deflate) an effect [96]. This type of misclassification was of less concern in this study, because the questionnaire sent to the cases and controls had prompted them to provide information on many risk factors and they would not have the knowledge or suspicion of the study hypothesis. Therefore, the possibility of distortion (differential recall) of the effects of main exposure variables due to the participants' awareness of the question under hypothesis in this study was unlikely. Since the collected data on the exposures was self-reported, differential misclassification, introduced by an interviewer or data coder was also less likely.

Given the above-mentioned, the observed patterns of risks for the main exposure variables (physical activity and body mass index) are more likely to reflect effects of non-differential rather than differential misclassification [95-97].

Selection Bias

Controls, if selected in an unbiased manner, should represent the source population which gave rise to the cases. Ideally, controls would estimate the expected exposure rate in cases had they not developed the disease [86; 98-99].

Since this was a population-based case-control study, it used a primary base (a geographically and temporally defined population experience). Both cases and controls were ascertained and sampled from the same primary base/source population of the Canadian population 20 years or older, during the same period of time. A control, if diagnosed with bladder cancer during the study period (1994-97), would become an eligible case [86; 98-99]. The randomly sampled controls, using population rosters such as Provincial health insurance databases, are more likely to provide representative (of the primary base) distributions of the main exposures. Random digit dialling was used only in two Provinces: Alberta and Newfoundland. Incomplete phone coverage, answering machines, multiple phone lines, and more than one eligible control per household could potentially lead to some selection bias. However, random digit dialling often overcomes the problem of unlisted numbers when telephone directories are used. The differences in sampling methods across the Provinces were accommodated in the analyses by adjusting the effect estimates for Province. Although the analytical adjustment for the selection-confounding marker (Province) would not completely remove the bias, it could have mitigated the extent and effect of selection bias.

Although complete case ascertainment in case-control studies with a primary base is a challenge, the cases in this study were identified and provided by the Provincial Cancer Registries, whose ascertainment rates are considered to be more or less complete (bladder cancer requires medical attention). This would facilitate cases' representativeness of the study base with respect to the exposures of interest.

Because the controls were selected and interviewed contemporaneously with the disease incidence, meaning that both cases and controls were sampled from the same risk-set (the controls were at risk at the time of cases' diagnoses), temporal coherence of cases and controls was achieved [86; 99].

Selection of controls must be implemented independently from the exposure under interest [99-100]. Since, in this study, controls had been selected randomly using population rosters (health insurance databases) and random digit dialling, it is unlikely that their selection was related to levels of recreational physical activity and relative weight. Hospital-based case-control studies are more prone to this type of bias when the condition for admission of controls is directly or indirectly related to the exposure of interest, leading to an artificial over or under-representation of the prevalence of exposure in the hospital control series.

The examination of the distributions of total physical activity scores and body mass index across the control subsets did not reveal any unusual patterns (Appendix B). These observed patterns of distribution conformed to expectations based on external knowledge

of empirical evidence [28]. The presence of expected patterns amongst the controls indicated that the control samples (men and women) had not been distorted due to some severe selection bias and/or loss of study participants (non-response).

Non-Response

Severe non-response bias would occur if the participation rates differed for cases and controls and the exposure distribution was different in non-responders relative to responders. It has often been documented that response rates amongst controls of low socioeconomic status (SES) are lower compared to those amongst controls of higher SES [101]. Due to this non-response, the control series may be distorted in a way that they would be artificially over-represented by subjects of higher SES relative to the study base (source of the general population that gave rise to the cases). Based on the well-known positive association between physical activity and socioeconomic status, one might suppose that the use of a control series with artificially over-represented subjects of higher SES (relative to the study base) would lead to the observation of an inverse association between physical activity and risk of bladder cancer (controls being more physically active than cases), even if there were no association between the two.

Since the proportions of non-response in this study were not low and they were similar for cases and controls, one could make a conjecture that the non-responders were of lower or similar SES (and physical activity levels) as responders, irrespective of the case-control status (non-response related only to the exposure but not the disease status). This would have maintained the proportionality in the case-control differences with respect to

SES, and consequently, the levels of recreational physical activity. Given the above-mentioned, one could take a modest comfort that this proportional loss would not have produced a substantial bias in the estimates of OR for physical activity. However, the non-response of this type could have reduced the study precision [86; 101]. The evidence that non-responders usually have lower or similar SES as responders, irrespective of the disease status, is supported by a large case-control study of breast cancer, based on the population of North American women of similar age, which showed that the non-responders had lower or similar SES than responders whether they were cases or controls [102].

Given similar non-response rates, severe non-response bias is a less likely scenario. It would be plausible only if the non-responders amongst cases were of higher SES/physical activity levels than the responders, and in controls, the non-responders were of lower SES/physical activity levels than the responders, or vice versa (combined response bias: non-response related to both the exposure such as SES/physical activity and the disease). Besides, SES should be a strong correlate of physical activity levels both in cases and controls.

Unfortunately, because of the lack of information, this thesis could not explore the distribution of socio-demographic characteristics of the non-responders in order to draw definitive conclusions regarding the extent and influence of non-response bias on the study results.

Residual Confounding

Since in this study cigarette smoking (pack-years) was a strong risk factor for bladder cancer (in other studies too) and it was categorized into quartiles, the observed associations might not have been completely adjusted for the effect of smoking. If subjects with lower physical activity levels were more likely to be heavy smokers, an incomplete adjustment for smoking could lead to an exaggeration of a true relative risk for the association between physical activity and bladder cancer risk in the direction of 'protective effect'.

Unknown and Missing Confounders

Some extent of confounding due to some unknown factors should not be excluded. In this case, it is hard to predict the direction of bias. There was missing information on certain risk factors of bladder cancer such as history of urinary tract infections, use of specific drugs (phenacetin-containing analgesics and cyclophosphamide), and family history. Fortunately, the prevalence of these factors in the Canadian general population is very low, and it is unlikely that their omission from the analyses produced major distortion in the effect estimates. Information on exogenous sex hormones (i.e., hormone replacement therapy and oral contraceptive use) was not available for this study.

Reverse Causality

The operation of reverse causality leading to curtailed levels of physical activity among the cases (i.e., the observed ORs in women with direction towards protective effect) can be ruled out. The examination of the bladder cancer cases with respect to the histological

types revealed that about 90% of them were diagnosed with transitional cell carcinoma/papilloma. According to the literature, the majority (about 75-80%) of all transitional cell tumours are superficial, papillary shaped, and have relatively benign nature (mostly non-invasive) with lower grade and stage at presentation. Most of the time, they are confined to the mucosal layer or the lamina propria. Most common early symptom in the presence of transitional cell tumour is painless or asymptomatic hematuria [24; 44-45; 103].

7. 4. Etiologic Considerations

Several biologic mechanisms are hypothesized to mediate an effect of increased levels of physical activity on risk of cancer. These protective mechanisms are concerned with decreased adiposity (body fat accumulation) [72-73], reduced exposure to endogenous sex hormones such as estrogens (estradiol, estriol, and estrone) and androgens (testosterone and dehydroepiandrosterone). The protective mechanisms of physical activity and/or lean body mass might also be related to decreased concentrations of insulin-like growth factors. Another mechanism suggests that regular moderate physical activity enhances immune function by increasing the secretion and circulation of lymphocytes, granulocytes, monocytes, and macrophages. All the above-mentioned factors are shown to be involved in carcinogenesis [27-28; 29-30; 32; 84]. Excess body weight (a surrogate of body fat mass) acts independently from physical activity to increase secretion and levels of plasma insulin, which is thought to increase the bioactivity of insulin-like growth factors and stimulate cell proliferation [27-30; 72-73]. Furthermore, a number of human and animal studies have demonstrated that certain endogenous sex hormones (testosterones,

progesterone, and estradiol) and insulin-like growth factor-I, enhanced the carcinogenesis of bladder cancer [11-12; 14-17; 19-20; 33].

The findings of this study suggesting that parous women (i.e., at least one live birth) had a decreased risk of bladder cancer compared to nulliparous women supports the earlier-stated hypothesis that parous women might be protected against bladder cancer through reduced exposure to estrogens.

7. 5. Study Advantages

This thesis is one of the first population-based studies to address the associations between self-reported recreational physical activity levels, relative weight, and risk for bladder cancer in men and women. The response rates across the cases and controls were relatively high.

Concurrent collection of information on the exposures from cases and controls would not violate the principle of comparable accuracy with respect to recall [82; 99].

The cases enrolled in this study were newly diagnosed. This has an advantage in terms of the interpretation of study results over the alternative of using prevalent cases. In prevalent cases, it is difficult to differentiate whether the observed association between the exposure and the disease reflects the effect of the exposure on the risk of disease or its prognosis/duration [82].

Compared to most other case-control studies of bladder cancer, this study had a large sample size which leads to an increased precision of the estimates. For example, in men, this study had an adequate power (at least 80%) to detect the true RR of 0.70 (or RR = 1.45) or more extreme.

Another advantage of this study is that we were able to control for most of the well-established bladder cancer risk factors which might have acted as confounders of the associations under investigation. In addition, stratification by gender and menopausal status allowed us the exploration of potential effect modification.

Even though the derived physical activity score was based on self-reported data, it incorporated all the necessary components (type, seasonality, session frequency, duration, and intensity) for relatively accurate classification of the study subjects with respect to their levels of recreational physical activity.

Since this study was population-based, given its internal validity, the results would be readily generalizable to the Canadian population.

7. 6. Study Limitations

Physical activity is a complex multidimensional construct, which is difficult to measure accurately. Therefore, there is a potential for measurement error. One of the main limitations of this study was the lack of information on recreational physical activity patterns pertaining to the etiologically relevant time-window(s). Since, the physical

activity self-reported data pertained to the 2 year time-period prior to the interviews, the relative ranks of the cases and controls with respect to physical activity levels and body mass index may have been different for this period as compared to those during the etiologically crucial time-period(s). Moreover, the data on physical activity and body mass index were collected only at one point in time. If the variables of interest were time-dependent, this would probably lead to some degree of misclassification (non-differential, differential, or both).

Another limitation was the lack of similar information on occupational physical activity patterns. If data on occupational physical activity had been available, the exploration of the association between physical activity and bladder cancer would have been more accurate and comprehensive.

The low prevalence of the exposure variable (recreational physical activity) in the study population may have led to a reduced precision of the estimates (wider 95% CIs) for the quartiles of recreational physical activity. In women, high levels or regular patterns of physical activity could have a slight protective effect on risk of bladder cancer through one or more hypothesized pathways (e.g., hormonal factors, insulin resistance, and/or immune function). However, this study in women had 60.0% or less power of detecting the magnitude of true OR = 0.70 or less extreme (see Table 1).

Although the questionnaire used for the data collection was highly structured, it was not validated in similar populations of men and women.

8. Summary of Findings

Based on the results of this study, levels of recreational physical activity and relative body weight (as measured by BMI) were not associated with risk of bladder cancer both in men and women. However, the study found that parous women (having at least one live birth) were at decreased risk of developing bladder cancer compared to nulliparous women.

Both in men and women, cigarette smoking was associated with increased risk of bladder cancer. Even though the number of pack-years of cigarette smoking in women was lower than in men, women had a higher smoking-related risk of bladder cancer than men. Women as well as men who had ever held any hazardous occupation(s) were at increased risk of bladder cancer compared to those who had never held such occupation(s). The occupation-related risk of bladder cancer was higher in women than in men. Coffee consumption was associated with a slightly elevated risk of bladder cancer in men, but not in women for whom there was no association between the two.

9. Future Research

Presently, it is not clear whether physical activity and body weight affect risk of bladder cancer. The results of this thesis pose questions that need to be addressed in future research. Namely, further research is needed to explore the associations between physical activity, relative weight, and bladder cancer. Ideally, this research would elucidate etiologic roles of physical activity, body fat distribution, and endogenous hormonal factors in the development of bladder cancer. Although plagued with feasibility problems, future studies should also consider multiple measurements of physical activity/energy expenditure and body fat distribution over time.

Some additional insights into gender-specific rate differences in bladder cancer could be gained by examining reproductive and hormonal factors in relation to bladder cancer.

It is desirable that the instruments for measuring levels of physical activity be tested for their validity and reproducibility to ensure more accurate data collection.

10. References

1. Brauers AB, Jakse G. Epidemiology and biology of human urinary bladder cancer. *J Cancer Res Clin Oncol* 2000; 126: 575-583.
2. Patton SE, Hall MG, Ozen H. Bladder cancer. *Curr Opin Oncol* 2002; 14: 265-272.
3. Johansson SL, Cohen SM. Epidemiology and etiology of bladder cancer. *Semin Surg Oncol* 1997; 13: 291-298.
4. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. *Int J Cancer* 1999; 80: 827-841.
5. Silverman DT, Hartge P, Morrison AS, et al. Epidemiology of bladder cancer. *Hematol Oncol Clin North Am* 1992; 6: 1-30.
6. National Cancer Institute of Canada. Canadian Cancer Statistics 2002. Toronto: The Institute; 2003.
7. Negri E, La Vecchia C. Epidemiology and prevention of bladder cancer. *Eur J Cancer Prevent* 2001; 10: 7-14.
8. Hartge P, Harvey EB, Linehan WM, et al. Unexplained excess risk of bladder cancer in men. *JNCI* 1990; 82: 1636-1640.
9. Anton-Culver H, Lee-Feldstein A, Taylor TH. The association of bladder cancer risk with ethnicity, gender, and smoking. *Ann Epidemiol* 1993; 3: 429-433.
10. Cantor KP, Lynch CF, Johnson D. Bladder cancer, parity, and age at first birth. *Cancer Causes and Control* 1992; 3: 57-62.
11. Yoshimoto J, Matsumura Y, Tanahashi T, et al. Effects of steroid sex hormones and adriamycin on human bladder cancer cells in culture. *Acta Medica Okayama* 1982; 36(1): 61-66.
12. Zhuang YH, Blauer M, Tammela T, et al. Immunodetection of androgen receptor in human urinary bladder cancer. *Histopathology* 1997; 30(6): 556-62.
13. Thomas DB. Steroid hormones and medications alter cancer risks. *Cancer* 1988; 62 (Suppl 8): 1755-67.
14. Reid LM, Leav I, Kwan PWL, Russel P, et al. Characterization of human, sex-steroid responsive transitional cell carcinoma maintained as a tumor line (R198) in athymic nude mice. *Cancer Res* 1984; 44: 4560-73.
15. Imaoka S, Yoneda Y, Sugimoto T, et al. Androgen regulation of CYP4B1 responsible for mutagenic activation of bladder carcinogens in the rat bladder: detection of CYP4B1

mRNA by competitive reverse transcription-polymerase chain reaction. *Cancer Lett* 2001; 166: 119-23.

16. Kono N, Tanahashi T, Suzawa N, et al. Effects of sex hormones on oncogenesis in rat urinary bladder by N-butyl-N-(4-hydroxylbutyl)-nitrosamine. *Int J Clin Pharmacol* 1977; 15: 101-105.
17. Horn EP, Tucker MA, Lambert G, et al. A study of gender-based cytochrome p4501A2 variability: a possible mechanism for the male excess of bladder cancer. *Cancer Epidemiol Biomarkers Prev* 1995; 4: 529-533.
18. Mostofi FK. Pathology of malignant tumors of the urinary bladder. The Biology and Clinical Management of Bladder Cancer (Cooper EH, Williams RE, eds). Philadelphia: Lippincott 1975.
19. Prianichnikova MB. Endogenous risk factors for bladder neoplasms in women. *Voprosy Onkologii* 1991; 37: 925-929.
20. Bertram JS, Graig AW. Specific induction of bladder cancer in mice by Butyl-(4-hydroxybutyl)-nitrosamine and the effects of hormonal modifications on the sex difference in response. *Europ J Cancer* 1972; 8: 587-594.
21. Gordon GB, Helzlsouer KJ, Comstock GW. Serum levels of dehydroepiandrosterone and its sulfate and the risk of developing bladder cancer. *Cancer Res* 1991; 51: 1366-1369.
22. Pelucchi C, La Vecchia C, Negri E, et al. Smoking and other risk factors for bladder cancer in women. *Prev Med* 2002; 35: 114-120.
23. Morrison AS, Cole P. Epidemiology of bladder cancer. *Urol Clin Nort Am* 1976; 3: 13-29.
24. Wai CY, Miller DS. Urinary bladder cancer. *Clin Obst Gynecol* 2002; 45: 844-854.
25. Ross RK, Jones PA, Yu MC. Bladder cancer epidemiology and pathogenesis. *Semin Oncol* 1996; 23: 536-545.
26. La Vecchia C, Airoldi L. Human bladder cancer: epidemiological, pathological, and mechanistic aspects. *International Agency for Research on Cancer (IARC)* 1999; 147: 139-157.
27. McTiernan A, Ulrich C, Slate S, et al. Physical activity and cancer etiology: associations and mechanisms. *Cancer Causes and Control* 1998; 9: 487-509.
28. IARC. *Weight control and physical activity, vol. 6. IARC Handbooks of Cancer Prevention.* International Agency for Research on Cancer: Lyon, France, 2001.
29. Thune I, Furberg AS. Physical activity and cancer risk: dose-response and cancer, all sites and site-specific. *Med Sci Sports Exerc* 2001; 33(6): S530-S550.

30. Kaaks R, Lukanova A. Effects of weight control and physical activity in cancer prevention. Role of endogenous hormone metabolism. *Ann N Y Acad Sci* 2002; 963: 268-281.
31. Yu H, Rohan T. Role of the insulin-like growth factor family in cancer development and progression. *JNCI* 2000; 92: 1472-89.
32. Hofmann-Goetz L. Influence of physical activity and exercise on innate immunity. *Nutr Rev* 1998; 56: S126-S130.
33. Zhao H, Grossman HB, Spitz MR, et al. Plasma levels of insulin-like growth factor-1 and binding protein-3, and their association with bladder cancer risk. *J Urol* 2003; 169: 714-717.
34. Torrisi R, Mezzetti M, Johansson H, et al. Time course of fenretidine-induced modulation of circulating insulin-like growth factor (IGF)-I, IGF-II and IGFBP-3 in a bladder cancer chemoprevention trial. *Int J Cancer* 2000; 87(4): 601-5.
35. Paffenbarger RS, Hyde RT, Wing AL. Physical activity and incidence of cancer in diverse populations: a preliminary report. *Am J Clin Nutr* 1987; 45: 312-7.
36. Severson RK, Nomura AM, Grove JS, et al. A prospective analysis of physical activity and cancer. *Am J Epidemiol* 1989; 130: 522-529.
37. Wannamethee SG, Shaper AG, Walker M. Physical activity and risk of cancer in middle-aged men. *Br J Cancer* 2001; 85(9): 1311-1316.
38. Tripathi A, Folsom AR, Anderson KE. Risk factors for urinary bladder carcinoma in postmenopausal women. *Cancer* 2002; 95: 2316-2323.
39. Whittemore AS, Paffenbarger RS, Anderson K, et al. Early precursors of urogenital cancers in former college men. *J Urol* 1984; 132: 1256-1261.
40. Dosemeci M, Hayes RB, Vetter R, Hoover RN, et al. Occupational physical activity, socioeconomic status, and risks of 15 cancer sites in Turkey. *Cancer Causes and Control* 1993; 4: 313-321.
41. Wolk A, Gridley G, Svensson M, et al. A prospective study of obesity and cancer risk (Sweden). *Cancer Causes and Control* 2001; 12: 13-21.
42. Adami HO, Hunter D, Trichopoulos D. Textbook of cancer epidemiology. Oxford, University Press 2002, New York.
43. International Union Against Cancer (UICC). In: Sobin LH, Wittekind C, eds. TNM classification of malignant tumors, 5th ed. New York: Wiley-Liss, 1997.
44. Aben KKH, Kiemeney ALM. Epidemiology of Bladder Cancer. *Eur Urol* 1999; 36(6): 1-13.

45. Kaye KW, Lange PH. Mode of presentation of invasive bladder cancer: reassessment of the problem. *J Urol* 1982; 128:31-33.
46. Puente D, Malats N, Cecchini L, et al. Gender-related differences in clinical and pathological characteristics and therapy of bladder cancer. *Eur Urol* 2003; 43: 53-62.
47. Mungan N, Kiemeney L, Van Dick J, et al. Gender differences in stage distribution of bladder cancer. *Urology* 2000; 55: 368-371.
48. Laor E, Schiffman Z, Reid RE, et al. Sex-dependent variations in natural history of transitional cell carcinoma of bladder. *Urology* 1984; 23(3): 237-239.
49. Hoke GP, Stone BA, Klein L, et al. The influence of gender on incidence and outcome of patients with bladder cancer in Harlem. *J Natl Med Assoc* 1999; 144-148.
50. Case RAM, Hosker ME, McDonald DM, et al. Tumors of urinary bladder in workmen engaged in manufacture and use of certain dye-stuff intermediates in British chemical industry: role of aniline, benzidine, alpha-naphtylamine, and beta-naphtylamine. *Br J Industr Med* 1954; 11: 75-104.
51. Siemiatycki J, Dewar R, Nadon L, et al. Occupational risk factors for bladder cancer: results from a case-control study in Montreal, Quebec, Canada. *Am J Epidemiol* 1994; 140: 1061-1080.
52. Silverman DT, Levin L, Hoover RN. Occupational risk of bladder cancer among white women in the United States. *Am J Epidemiol* 1990; 132: 453-461.
53. Steinmaus CM, Nunez S, Smith AH. Diet and bladder cancer: A meta-analysis of six dietary variables. *Am J Epidemiol* 2000; 151: 693-702.
54. Viscoli CM, Lachs MS, Horwitz RI. Bladder cancer and coffee drinking: a summary of case-control research. *Lancet* 1993; 341: 1432-1437.
55. Sala M, Cordier S, Chang-Claude J, et al. Coffee consumption and bladder cancer in nonsmokers : a pooled analysis of case-control studies in European countries. *Cancer Causes and Control* 2000; 11: 925-931.
56. Zeegers MP, Tan FES, Verhagen A, et al. Elevated risk of cancer of the urinary tract for alcohol drinkers: a meta-analysis. *Cancer Causes and Control* 1999; 10: 445-451.
57. Pelucchi C, Negri E, Franceschi S, et al. Alcohol drinking and bladder cancer. *J Clin Epidemiol* 2002; 55: 637-641.
58. Villanueva CM, Fernandez F, Malats N, et al. Meta-analysis of studies on individual consumption of chlorinated drinking water and bladder cancer. *J Epidemiol Community Health* 2003; 57: 166-173.

59. King WD, Marrett LD. Case-control study of bladder cancer and chlorination by-products in treated water (Ontario, Canada). *Cancer Causes and Control* 1996; 7: 796-604.
60. Cantor KP, Lynch F, Hildesheim ME, et al. Drinking water source and chlorination byproducts. I. Risk of bladder cancer. *Epidemiology* 1998; 9: 21-28.
61. Wu MM, Kuo TL, Hwang YH, et al. Dose-response relation between arsenic concentration in well water and mortality from cancers and vascular diseases. *Am J Epidemiol* 1989; 130; 1123-1132.
62. Chen CJ, Chuang YC, Lin TM, et al. Malignant neoplasms among residents of a blackfoot disease-endemic area in Taiwan: high arsenic artesian well water and cancers. *Cancer Res* 1985; 45: 5895-5899.
63. Michaud DS, Spiegelman D, Clinton SK, et al. Fluid intake and the risk of bladder cancer in men. *N Engl J Med* 1999; 340: 1390-1397.
64. Zeegers MP, Dorant E, Goldbohm RA, et al. Are coffee, tea, and total fluid consumption associated with bladder cancer risk? Results from the Netherlands Cohort study. *Cancer Causes and Control* 2001; 12: 231-238.
65. Slattery ML, West DW, Robinson LM. Fluid intake and bladder cancer in Utah. *Int J Cancer* 1988; 42: 17-22.
66. Pederson-Bjergaard J, Ersboll J, Hansen VL, et al. Carcinoma of the urinary bladder after treatment with cyclophosphamide for non-Hodgkin's lymphoma. N Engl J Med 1988; 318: 1028-1032.
67. Kantor AF, Hartge P, Hoover RN, et al. Familial and environmental interactions in bladder cancer risk. *Int J Cancer* 1985; 35: 703-706.
68. National Academy of Sciences: health risks of exposure to low levels of ionizing radiation (BEIR V). Washington, National Academy Press, 1990.
69. Pols MA, Peeters PHM, Kemper HC, et al. Methodological aspects of physical activity assessment in epidemiological studies. *Eur J Epidmiol* 1998; 14: 63-70.
70. Oliveria SA, Christos PJ. The epidemiology of physical activity and cancer. *Ann N Y Acad Sci* 1997; 833: 79-90.
71. Prentice AM. Beyond body mass index. *Obes Rev* 2001; 2(3): 141-147.
72. Fogelholm M, Vainio H. Weight control, physical activity and cancer – strong links. *Obes Rev* 2002; 3: 1-3.
73. Bianchini F, Kaaks R, Vainio H. Weight control and physical activity in cancer prevention. *Obes Rev* 2002; 3: 5-8.

74. Green A, Beral V, Moser K. Mortality in women in relation to their childbearing history. *BMJ* 1988; 297: 391-401.

75. Plesko I, Preston-Martin S, Day NE, et al. Parity and cancer risk in Slovakia. *Int J Cancer* 1985; 36: 529-533.

76. Miller AB, Barklay THC, Choi MG, et al. A study of cancer, parity, and age at first pregnancy. *J Chron Dis* 1979; 33: 595-605.

77. Kabat GC, Dieck GS, Wynder EL. Bladder cancer in nonsmokers. *Cancer* 1986; 57: 362-367.

78. Moller H, Mellemgaard A, Lindvig K, et al. Obesity and cancer risk: a Danish record-linkage study. *Eur J Cancer* 1994; 30A: 344-450.

79. Brownson RC, Chang JC, Davis JR, et al. Physical activity on the job and cancer in Missouri. *J Public Health* 1991; 81: 639-642.

80. Johnson KC, Mao Y, Argo J, et al. The National Enhanced Cancer Surveillance System: A case-control approach to environment-related cancer surveillance in Canada. *Environmetrics* 1998; 9: 495-504.

81. World Health Organization. International Classification of Diseases for Oncology. Third edition. Geneva, Switzerland: World Health Organization, 2000.

82. Kelsey JL, Whittemore AS, Evans AS, et al. Methods in observational epidemiology. New York: Oxford University Press, 1996.

83. Statistics Canada. *Standard Occupational Classification.* Ottawa: Statistics Canada, Standards Division, 1980.

84. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exer* 1993; 25: 71-80.

85. Friedenreich CM, Bryant HE, Courneya KS, et al. Case-control study of lifetime physical activity and breast cancer risk. *Am J Epidemiol* 2001; 154: 336-347.

86. Savitz DA. Interpreting epidemiologic evidence: strategies for study design and analysis. New York: Oxford University Press; 2003.

87. Hosmer DW, Lemeshow S. Applied logistic regression. New York: Wiley; 1989.

88. Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* 1996; 49: 907-916.

89. Kelsey J. Breast cancer: magnitude of the problem and descriptive epidemiology. *Epidemiol Rev* 1993; 15: 256-263.

90. Helmrich SP, Shapiro S, Rosenberg L, et al. Risk factors for breast cancer. *Am J Epidemiol* 1983; 117: 35-45.

91. Kuper H, Cramer DW, Titus-Ernstoff L. Risk of ovarian cancer in the United States in relation to anthropometric measures: does the association depend on menopausal status? *Cancer Causes and Control* 2002; 13: 455-463.

92. SAS Institute Inc. SAS Guide for personal computers, version 8.0. SAS Institute Inc., Cary, NC, 2000.

93. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of event per variable in logistic regression analysis. *J Clin Epidemiol* 1996; 49 (12): 1373-1379.

94. Copeland KT, Checkoway H, McMichael AJ, et al. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977; 105: 488-495.

95. Dosemeci M, Wacholder S, Lubin J. Does non-differential misclassification of exposure always bias a true effect toward the null value? *Am J Epidemiol* 1990; 132: 746-748.

96. Rothman KJ. Modern epidemiology. Boston: Little, Brown and Company, 1986.

97. Birkett NJ. Effect of nondifferential misclassification on estimates of odds ratios with multiple levels of exposure. *Am J Epidemiol* 1992; 136 (3): 356-362.

98. Miettinen OS. The "case-control" study: valid selection of subjects. *J Chronic Dis* 1985; 38: 543-548.

99. Wacholder S, McLaughlin JK, Silverman DT, et al. Selection of controls in case-control studies. *Am J Epidemiol* 1992; 135: 1019-1028.

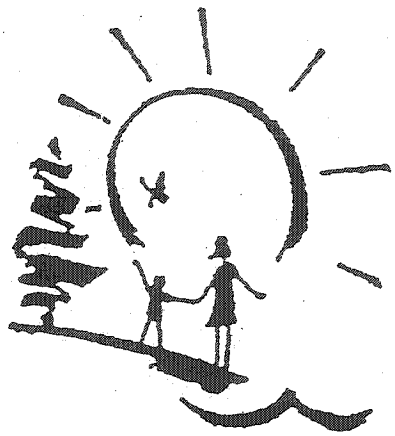
100. Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet* 2002; 359: 431-434.

101. Richiardi L, Boffetta P, Merletti F. Analysis of nonresponse bias in a population-based case-control study on lung cancer. *J Clin Epidemiol* 2002; 55: 1033-1040.

102. Moorman PG, Newman B, Millian RC, et al. Participation rates in a case-control study: the impact of age, race, and race of interviewer. *Ann Epidemiol* 1999; 9: 188-195.

103. Kiemeney LALM, Witjes JA, Verbeek ALM, et al. The clinical epidemiology of superficial bladder cancer. *Br J Cancer* 1993; 67: 806-812.

Appendix A – NECSS Questionnaire



Environmental Health Survey

(Confidential when completed)

This form asks a variety of questions about you and your environment, which may affect or be related to your health. The information you provide will help Canadians to understand more about preventing disease.

Please complete each question as best you can even if you are not sure of your answer.

You do not need to fill in the entire questionnaire all at once. You may wish to take a brief rest in the middle.

If you have any questions about the survey or would like help filling it out, please call

_____

at _____.

Please return this questionnaire by

_____. Thank you for your time.

Guide to filling in this Questionnaire:

Please choose answers by marking a circle ○ — ✓

or printing in the boxes. ☐ — ☐

☐ — April

☐ **Si vous préférez répondre en français, veuillez cocher la case et renvoyer le questionnaire dans l'enveloppe adressée ci-jointe.**

GENERAL INFORMATION

141

1. Today's date — *Month* ____ *Day* ____ *Year* 19____

2. **Is anyone helping you** (the person whose name appears on the front of the questionnaire) to complete this questionnaire?
   - o No
   - o Yes → o Spouse
   - o Other -- Please specify: ____

3. When were you born?
   Month ____ *Day* ____ *Year* 19____

4. Are you
   - o Female
   - o Male

5. To which **ethnic or cultural group(s)** did **your ancestors** belong?

   Mark or specify as many as applicable.
   - o French
   - o English
   - o German
   - o Scottish
   - o Italian
   - o Irish
   - o Ukrainian
   - o Chinese
   - o Dutch
   - o Jewish
   - o Polish
   - o Black
   - o Aboriginal
   - o Métis
   - o Inuit
   - o Other ethnic or cultural group(s) -- Please specify: ____

   Examples of other ethnic or cultural groups are: Portuguese, Greek, Indian, Pakistani, Vietnamese, Japanese, Lebanese, Haitian, etc.

6. What is your **marital status**?
   - o Single
   - o Married
   - o Common law
   - o Widowed
   - o Divorced/Separated
   - o Other

7. What is the **highest grade** (or year) of high school or elementary school that you have completed?
   Grade ____ o Never attended school

8. How many **years of post-secondary** school have you completed?
   ____ years o none

9. Have you smoked **at least 100 cigarettes** in your entire life?
   - o No → Go to 10
   - o Yes

   About how old were you when you **first started** smoking cigarettes? ____ years old

   About **how many years** in total did you smoke? ____ years

   Of the entire time you smoked, how many **cigarettes**, on the average, did you smoke **per day**? ____ per day

   Do you smoke cigarettes now?

   - o No → How old were you when you **stopped** smoking? ____ years old
   - o Yes → On the average, about how many **cigarettes a day** do you smoke now? ____ per day

10. Have you ever smoked a **pipe or cigars** regularly?
    - o No → Go to 11
    - o Yes

    For how many years? ____ years
    About how many pipes or cigars? ____ per day or ____ per week

11. Have you ever used **chewing tobacco** regularly?
    - o No → Go to 12
    - o Yes

    For how many years? ____ years
    About how many plugs? ____ per day or ____ per week

12. How **tall** are you?
    ____ feet ____ inches or ____ centimetres

13. How much did you **weigh** about 2 years ago?
    ____ pounds or ____ kilograms

14. What is the **most you have ever weighed**? (Women should not include pregnancy.)
    ____ pounds or ____ kilograms

RESIDENTIAL HISTORY

15. Please list each of the places **in Canada** you have lived for at least **1 year.** Start with the most recent residence and follow back to your childhood. (If you cannot remember exact details, provide your best recollection, for example, nearest cross-street or intersection.)

	TIME PERIOD	ADDRESS			
	First Year *Last Year*	Street and Number or Lot, Concession and Township **(If you don't remember the exact address, give the nearest cross-street or intersection.)**	City, Town, or Municipality	County or District, if rural	Province
	Example: 19 8 9 to 19 9 3	*97 Greargate Rd.*	*Brandon*		*Manitoba*
1	19 to 19				
2	19 to 19				
3	19 to 19				
4	19 to 19				
5	19 to 19				
6	19 to 19				
7	19 to 19				
8	19 to 19				
9	19 to 19				
10	19 to 19				
11	19 to 19				
12	19 to 19				

ADDRESS / Postal Code	Main source of **drinking water** (Town/city water, Dug well, Drilled well, Bottled, Other, Don't know)	Primary types of **home heating** (Mark those which apply.) (Oil, Natural gas/propane, Electric, Wood, Coal, Other, Don't know)	Were you aware of **dusts or odours from industry** while living at this residence? (Never or rarely, At least once a month, At least once a week, At least once a day)	How many **regular smokers** usually lived in this home with you? (None, 1, 2, 3 or more, Don't know)	143
R7A 8A7	✓ ○ ○ ○ ○ ○	○ ✓ ○ ✓ ○ ○ ○	○ ✓ ○ ○	✓ ○ ○ ○ ○	
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	1
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	2
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	3
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	4
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	5
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	6
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	7
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	8
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	9
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	10
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	11
	○ ○ ○ ○ ○ ○	○ ○ ○ ○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	12

EMPLOYMENT HISTORY

16. Please tell us about **each job or occupation you had for at least 12 months in Canada or elsewhere.** Include seasonal work, part-time, etc., if you worked the equivalent of 12 months or more. Begin with your most recent job and continue back to your first job. Please estimate the time period if you cannot remember exact years. (Even if you have retired, we still require the information.) **If you have never been employed, check here** ☐ and continue to Question 17.

	TIME PERIOD *First Year* *Last Year*	Type of Industry, Business, or Service and Company Name	Main Job Duties
	Example: 19 89 to 19 94	*Oil industry, Bluestar Oil Company*	*Process powerhouse and sulphur plant operations.*
1	19 ☐☐ to 19 ☐☐		
2	19 ☐☐ to 19 ☐☐		
3	19 ☐☐ to 19 ☐☐		
4	19 ☐☐ to 19 ☐☐		
5	19 ☐☐ to 19 ☐☐		
6	19 ☐☐ to 19 ☐☐		
7	19 ☐☐ to 19 ☐☐		
8	19 ☐☐ to 19 ☐☐		
9	19 ☐☐ to 19 ☐☐		
10	19 ☐☐ to 19 ☐☐		

Job Location(s) *City/town and province*	Job Title	Status Full time / Part time / Seasonal / Other	At this work-place, were you aware of **dusts or odours** from industry? Never or rarely / At least once a month / At least once a week / At least once a day	About how many people **smoked regularly** in your **immediate work area?** None / 1 or 2 / 3 to 5 / 6 or more / Don't know	
Rocky Mountain House, Alberta	*Gas plant operator*	✓ ○ ○ ○	○ ○ ○ ✓	○ ○ ✓ ○ ○	
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	1
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	2
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	3
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	4
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	5
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	6
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	7
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	8
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	9
		○ ○ ○ ○	○ ○ ○ ○	○ ○ ○ ○ ○	10

17. Have you ever worked with any of the following **for more than one year?**

	Never	Don't know	At work	At home	How many years in total?
Asbestos	o	o	o	o	
Arsenic salts	o	o	o	o	
Chromium salts	o	o	o	o	
Cadmium salts	o	o	o	o	
Coal tar, soot, pitch, creosote, asphalt	o	o	o	o	
Mineral, cutting or lubricating oil	o	o	o	o	
Benzidine	o	o	o	o	
Benzene	o	o	o	o	
Isopropyl oil	o	o	o	o	
Dyestuffs	o	o	o	o	
Vinyl chloride	o	o	o	o	
Pesticides	o	o	o	o	
Herbicides	o	o	o	o	
Mustard gas	o	o	o	o	
Radiation sources	o	o	o	o	
Welding	o	o	o	o	
Wood dust	o	o	o	o	

You are more than halfway through the questionnaire. This is a good place to take a short break, if you wish.

DIET INFORMATION

18. During the past 20 years have you ever taken any of the following vitamin or mineral supplements?

Vitamin and Mineral type	How often?			For how many years in total?					
	No	Yes, but not regularly	Yes, fairly regularly	Less than 1 year	1 to 2 years	3 to 5 years	6 to 9 years	10 to 19 years	20 + years
Multiple vitamins	o	o	o	o	o	o	o	o	o
Vitamin A	o	o	o	o	o	o	o	o	o
Beta-carotene	o	o	o	o	o	o	o	o	o
B-complex vitamins	o	o	o	o	o	o	o	o	o
Vitamin C	o	o	o	o	o	o	o	o	o
Vitamin E	o	o	o	o	o	o	o	o	o
Calcium	o	o	o	o	o	o	o	o	o
Iron	o	o	o	o	o	o	o	o	o
Zinc	o	o	o	o	o	o	o	o	o
Selenium	o	o	o	o	o	o	o	o	o

19. This section asks about your eating habits **about two years ago.** Thinking back to that time, we ask you to mark the column that best describes how often, on average, you ate or drank the amount specified of each of the following foods and beverages.

	Never or less than 1 per month	1-3 per month	1 per week	2-4 per week	5-6 per week	1 per day	2-3 per day	4-5 per day	6+ per day
BEVERAGES MADE WITH WATER									
Coffee (1 cup)	O	O	O	O	O	O	O	O	O
Tea (1 cup)	O	O	O	O	O	O	O	O	O
Orange or grapefruit juice from frozen concentrate (4 oz/115 ml glass)	O	O	O	O	O	O	O	O	O
Other juices or drinks from frozen concentrate (4 oz/115 ml glass)	O	O	O	O	O	O	O	O	O
Drinks from powdered drink crystals (4 oz/115 ml glass)	O	O	O	O	O	O	O	O	O
Tap water (8 oz/230 ml glass)	O	O	O	O	O	O	O	O	O
Bottled water (8 oz/230 ml glass)	O	O	O	O	O	O	O	O	O
OTHER BEVERAGES									
Whole milk (8 oz/230 ml glass)	O	O	O	O	O	O	O	O	O
2% milk (8 oz/230 ml glass)	O	O	O	O	O	O	O	O	O
1% milk (8 oz/230 ml glass)	O	O	O	O	O	O	O	O	O
Skim milk (8 oz/230 ml glass)	O	O	O	O	O	O	O	O	O
Orange or grapefruit juice, fresh, bottled or canned (4 oz/115 ml glass)	O	O	O	O	O	O	O	O	O
Other juices or drinks, fresh, bottled or canned (4 oz/115 ml glass)	O	O	O	O	O	O	O	O	O
Tomato or vegetable juices (4 oz/115 ml glass)	O	O	O	O	O	O	O	O	O
Soft drinks (1 glass/bottle/can)	O	O	O	O	O	O	O	O	O
Beer (1 bottle/can)	O	O	O	O	O	O	O	O	O
Wine (1 glass)	O	O	O	O	O	O	O	O	O
Liquor (1 drink or shot)	O	O	O	O	O	O	O	O	O

(continued)

19. (continued) Please mark the column that best describes how often, on average, you ate the amount specified of these foods about two years ago.

	Never or less than 1 per month	1-3 per month	1 per week	2-4 per week	5-6 per week	1 per day	2-3 per day	4-5 per day	6 + per day
FRUIT									
Apples or pears (1)	o	o	o	o	o	o	o	o	o
Oranges (1)	o	o	o	o	o	o	o	o	o
Bananas (1)	o	o	o	o	o	o	o	o	o
Cantaloupe (¼ melon)	o	o	o	o	o	o	o	o	o
Other fruit, fresh or canned (1 piece or ½ cup)	o	o	o	o	o	o	o	o	o
VEGETABLES									
Tomatoes (1 or ½ cup)	o	o	o	o	o	o	o	o	o
Carrots (1 whole or ½ cup)	o	o	o	o	o	o	o	o	o
Broccoli (½ cup)	o	o	o	o	o	o	o	o	o
Cabbage, cauliflower, brussels sprouts (½ cup)	o	o	o	o	o	o	o	o	o
Spinach or other greens (1 serving)	o	o	o	o	o	o	o	o	o
Yellow (winter) squash (½ cup)	o	o	o	o	o	o	o	o	o
Any other vegetable including green beans, corn and peas (½ cup)	o	o	o	o	o	o	o	o	o
Soups with vegetables (1 cup)	o	o	o	o	o	o	o	o	o
Potatoes: baked, boiled (1) or mashed (1 cup)	o	o	o	o	o	o	o	o	o
French fries or fried potatoes (½ cup)	o	o	o	o	o	o	o	o	o
Sweet potatoes (1 or ½ cup)	o	o	o	o	o	o	o	o	o
Tofu or soybeans (3-4 oz/115 ml)	o	o	o	o	o	o	o	o	o
Baked beans or lentils (½ cup)	o	o	o	o	o	o	o	o	o
BREADS AND CEREALS									
Bran or granola cereals, shredded wheat (1 cup)	o	o	o	o	o	o	o	o	o
Other cold cereals (1 cup)	o	o	o	o	o	o	o	o	o
Cooked cereals (1 cup)	o	o	o	o	o	o	o	o	o
White bread (1 slice) or rolls (1)	o	o	o	o	o	o	o	o	o
Dark or whole grain bread (1 slice) or rolls (1)	o	o	o	o	o	o	o	o	o
Rice (1 cup)	o	o	o	o	o	o	o	o	o
Macaroni, spaghetti or noodles (1 cup)	o	o	o	o	o	o	o	o	o

(continued)

19. (continued) Please mark the column that best describes how often, on average, you ate the amount specified of these foods **about two years ago.**

	Never or less than 1 per month	1-3 per month	1 per week	2-4 per week	5-6 per week	1 per day	2-3 per day	4-5 per day	6 + per day
MEAT, POULTRY, FISH, EGGS & CHEESE									
Chicken or turkey (4 oz/115 ml)	O	O	O	O	O	O	O	O	O
Beef, pork or lamb as a main dish (steak, roast, ham) (4 oz/115 ml)	O	O	O	O	O	O	O	O	O
Beef, pork or lamb as a mixed dish (stew or casserole, pasta dish) (4 oz/115 ml)	O	O	O	O	O	O	O	O	O
Hamburger (1)	O	O	O	O	O	O	O	O	O
Hot dogs (1)	O	O	O	O	O	O	O	O	O
Luncheon meats (salami, bologna) (1 piece or slice)	O	O	O	O	O	O	O	O	O
Smoked meat or corned beef (1 piece or slice)	O	O	O	O	O	O	O	O	O
Bacon (1 slice)	O	O	O	O	O	O	O	O	O
Sausage (1)	O	O	O	O	O	O	O	O	O
Liver (4 oz/115 ml)	O	O	O	O	O	O	O	O	O
Fish, fresh, frozen or canned (4 oz/115 ml)	O	O	O	O	O	O	O	O	O
Fish, smoked, salted or dried (4 oz/115 ml)	O	O	O	O	O	O	O	O	O
Eggs (1)	O	O	O	O	O	O	O	O	O
Cheese other than cottage cheese (1 slice or 1 oz)	O	O	O	O	O	O	O	O	O
SWEETS									
Cake (1 slice)	O	O	O	O	O	O	O	O	O
Cookies (1)	O	O	O	O	O	O	O	O	O
Doughnuts, pastry (1)	O	O	O	O	O	O	O	O	O
Pies (1 slice)	O	O	O	O	O	O	O	O	O
Ice cream (½ cup)	O	O	O	O	O	O	O	O	O
Chocolate (1 small bar or 1 oz)	O	O	O	O	O	O	O	O	O
MISCELLANEOUS									
Potato chips (small bag or 45g)	O	O	O	O	O	O	O	O	O
Peanut butter (1 tbsp)	O	O	O	O	O	O	O	O	O
Nuts (1 oz/30g)	O	O	O	O	O	O	O	O	O
Margarine on bread or vegetables (1 pat or tsp)	O	O	O	O	O	O	O	O	O
Butter on bread or vegetables (1 pat or tsp)	O	O	O	O	O	O	O	O	O
Mayonnaise or salad dressing on bread or in salads (1 tbsp)	O	O	O	O	O	O	O	O	O

20. About two years ago:

	Seldom or Never	Sometimes	Often or Always
How often did you **add salt** to your food?	o	o	o
How often did you **add pepper** to your food?	o	o	o
How often did you have **onions or garlic** in your food?	o	o	o
How often did you **eat the skin on chicken**?	o	o	o
How often did you **eat the fat on meat**?	o	o	o

21. What kinds of **fat** did **you usually use** in cooking about 2 years ago?
Mark only 1 or 2.

- o Block or stick margarine
- o Soft tub margarine
- o Low-calorie margarine
- o Shortening
- o Butter
- o Oil
- o Lard, baconfat, fatback
- o Non-stick spray or no fat
- o Don't know or don't cook

22. What kinds of **fat** did **you usually put** on bread, potatoes and vegetables about 2 years ago?
Mark only 1 or 2.

- o Block or stick margarine
- o Soft tub margarine
- o Low-calorie margarine
- o Butter
- o Cream cheese
- o Didn't add fat

23. Summary Questions

About 2 years ago:

	Never or less than 1 per month	Less than 1 per week	1 to 2 per week	3 to 4 per week	5 to 6 per week	1 per day	1½ per day	2 per day	3 per day	4 + per day
How often did you use **fat or oil in cooking?**	o	o	o	o	o	o	o	o	o	o
Not counting salad or potatoes, **how many servings of vegetables** did you eat?	o	o	o	o	o	o	o	o	o	o
Not counting juices, how many **servings of fruit** did you eat?	o	o	o	o	o	o	o	o	o	o
How often did you eat **fried food** from a **restaurant or take-out?** (for example, french fries, fried chicken, fried fish)	o	o	o	o	o	o	o	o	o	o

24. We have a few questions about your **usual eating habits 20 years ago.** What you have just told us about the different places you have lived and worked might help in remembering back to your eating habits at that time.

For each of the following foods and beverages, please indicate **whether you usually ate or drank more or less 20 years ago** than you did about 2 years ago. Please mark the appropriate column for each food or beverage.

Compared to 2 years ago, 20 years ago I used to consume:	Much less	Some-what less	About the same amount	Some-what more	Much more
Beef, pork or lamb	o	o	o	o	o
Chicken or fish	o	o	o	o	o
Milk	o	o	o	o	o
Vegetables	o	o	o	o	o
Fruit	o	o	o	o	o
Bread	o	o	o	o	o
Margarine	o	o	o	o	o
Butter	o	o	o	o	o
Fried foods	o	o	o	o	o
Sweets	o	o	o	o	o
Tea	o	o	o	o	o
Coffee	o	o	o	o	o
Soft Drinks	o	o	o	o	o
Alcohol	o	o	o	o	o

25. About **how many times** have you gone on a **diet to lose weight** during your adult life?

- o Never
- o 1 to 2 times
- o 3 to 5 times
- o 6 to 8 times
- o 9 to 11 times
- o 12 or more times

PHYSICAL ACTIVITY

26. About 2 years ago, how often did you do the following activities, on average?

	Never	Which Seasons? Spring	Summer	Fall	Winter	How often? Less than 1 per month	1-3 per month	1-2 per week	3-6 per week	Every day	Time per session Less than 15 minutes	15-30 minutes	31-60 minutes	60+ minutes
Walking for exercise	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Jogging or running	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Gardening or yard work	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Home exercises or exercise class	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Golf	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Racquet sports (tennis, squash, etc.)	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Bowling or curling	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Swimming or water exercises	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Skiing or skating	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Bicycling	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Social dancing	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Other strenuous exercise	o	o	o	o	o	o	o	o	o	o	o	o	o	o

OTHER GENERAL INFORMATION

27. What was the **approximate total income for all household members** from all sources, **before income taxes,** in an average year during the last 5 years?

- o less than $10,000
- o $10,000 - $19,999
- o $20,000 - $29,999
- o $30,000 - $49,999
- o $50,000 - $99,999
- o greater than $100,000
- o prefer not to answer

28. How many members (adults and children) are there in your household in total?

[ ] persons

WOMEN:
Please continue to the next page.

MEN:
You have now completed the questionnaire.

Please take a moment to fill in any questions you may have missed.

THANK YOU VERY MUCH for taking the time to fill out this questionnaire. Your participation is sincerely appreciated.

Please return this completed questionnaire in the self-addressed envelope.

QUESTIONS FOR FEMALES ONLY

MENSTRUATION

29. How old were you when you had your first menstrual period?
    ☐ years old

    o Don't remember
    o Haven't menstruated → Go to 33

30. Between the ages of 10 and 30, did your menstrual periods tend to occur regularly or irregularly (menstrual cycles varied by more than 10 days in length)? Please exclude any time when you were pregnant or using birth control pills.
    o Regularly o Irregularly

31. **How old** were you when you had your **last menstrual period**?
    ☐ years old

    o Still menstruate → Go to 33

32. How did your **menstrual periods stop**?
    o Naturally--that is, as part of the change of life
    o As a result of a hysterectomy (removal of womb)
    o Following radiation
    o Other -- Please specify:
    ☐

33. Have you had an operation to remove **BOTH your ovaries**?
    o No
    o Yes → At what age? If they were removed on two separate occasions, record when your second ovary was removed.
    At age ☐ years

34. Do you have **mammograms** (x-rays of the breast) performed on a **routine** basis (every two years)?
    o No
    o Yes → First mammogram at age ☐ years

PREGNANCIES

35. Have you ever been pregnant?
    o No → Go to bottom of page.
    o Yes

36. **How many times have you been pregnant?** Include live births, stillbirths, miscarriages, abortions and ectopic (tubal) pregnancies.
    ☐ times

37. How old were you at the end of your first **pregnancy**?
    ☐ years old

38. How many **of your pregnancies** were live births?
    ☐ live births o none

39. How old were you at the end of your first pregnancy which lasted 5 months or more?
    ☐ years old

40. For how many **months** in total did you **breastfeed**? (Add the number of months that you breastfed after each birth to give the total number of months.)
    ☐ months o never breastfed

Please take a moment to fill in any questions you may have missed.

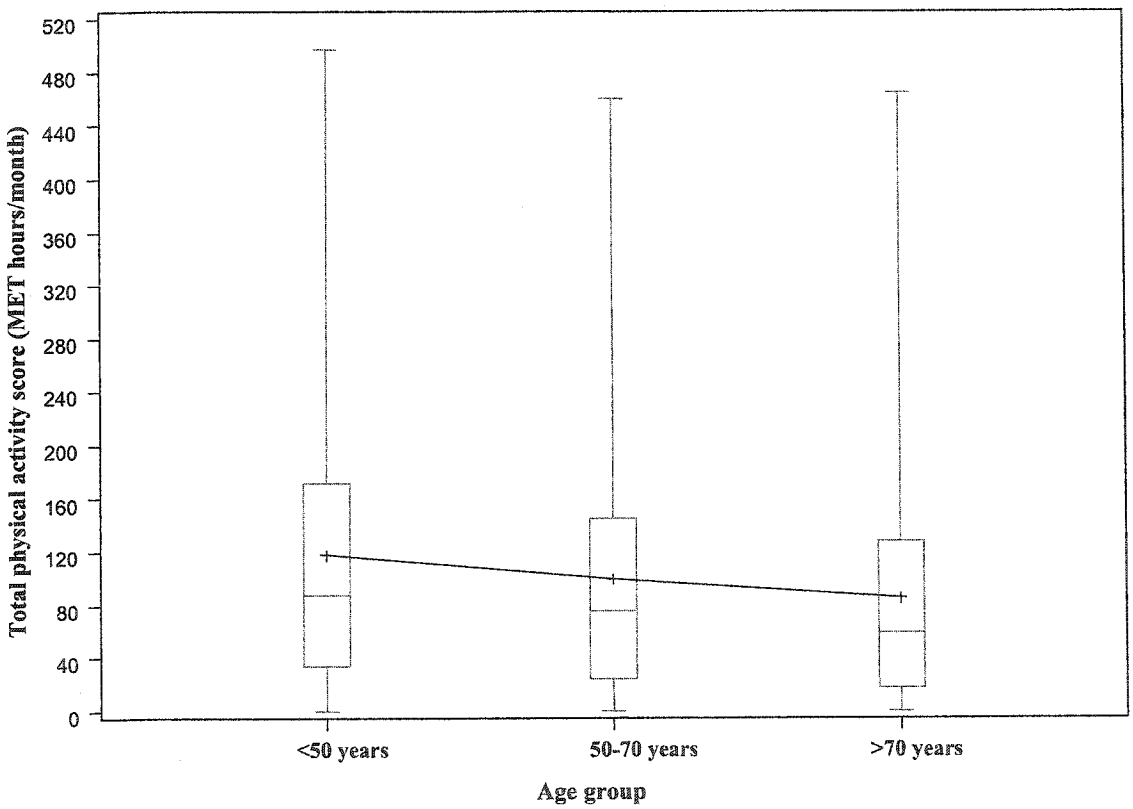
THANK YOU VERY MUCH for taking the time to fill out this questionnaire. Your participation is sincerely appreciated.

Please return this completed questionnaire in the self-addressed envelope.

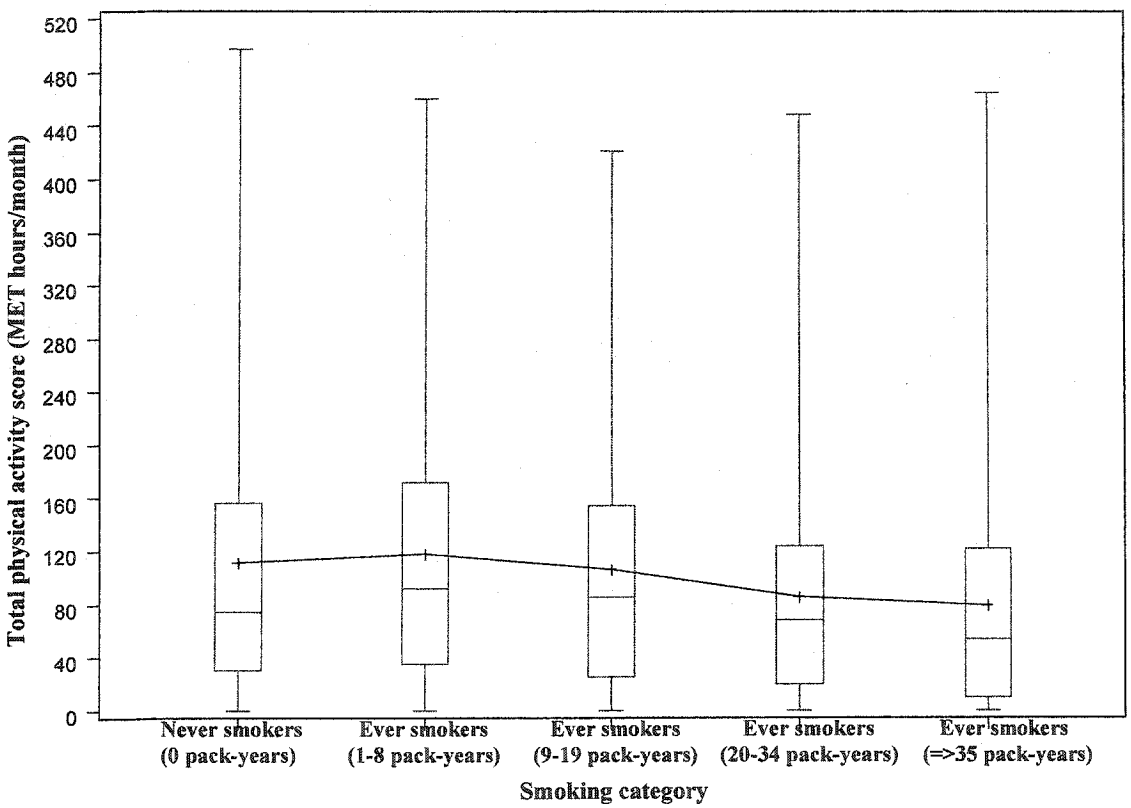
Appendix B – Distribution of Total Physical Activity Score and Body Mass Index in Controls

Appendix B1 – Male Controls

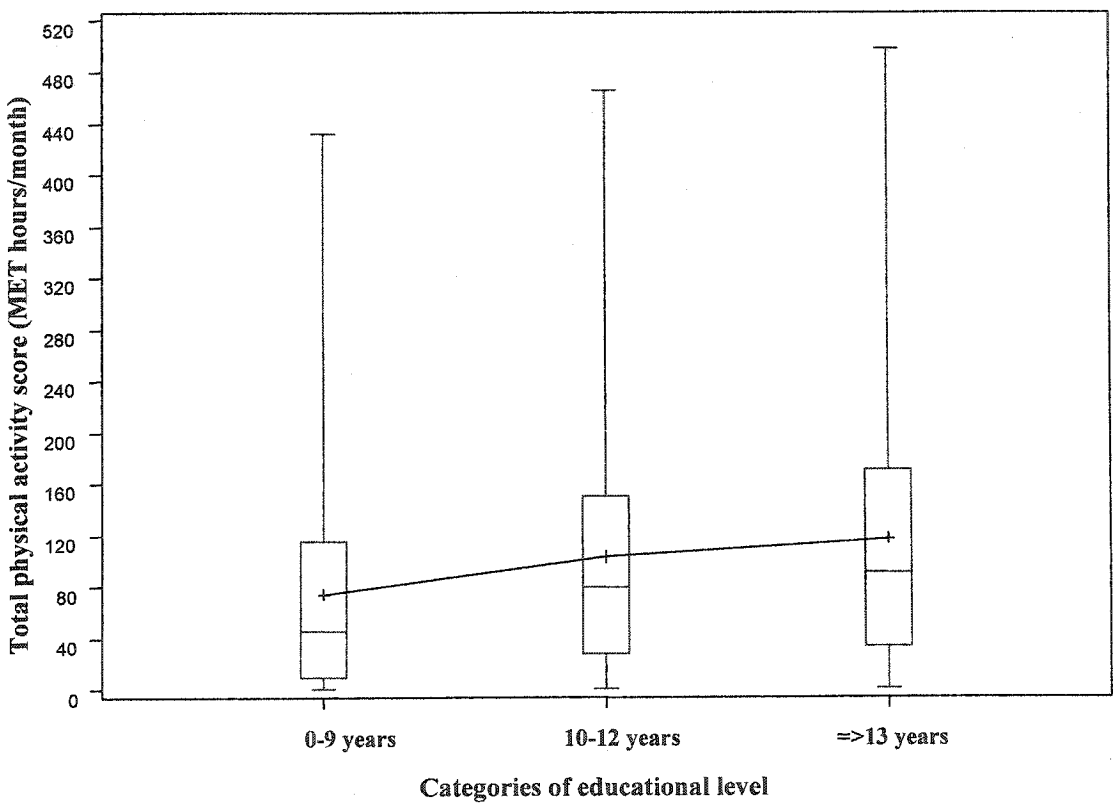
Box Plot B1-1 Total physical activity score by age group in male controls



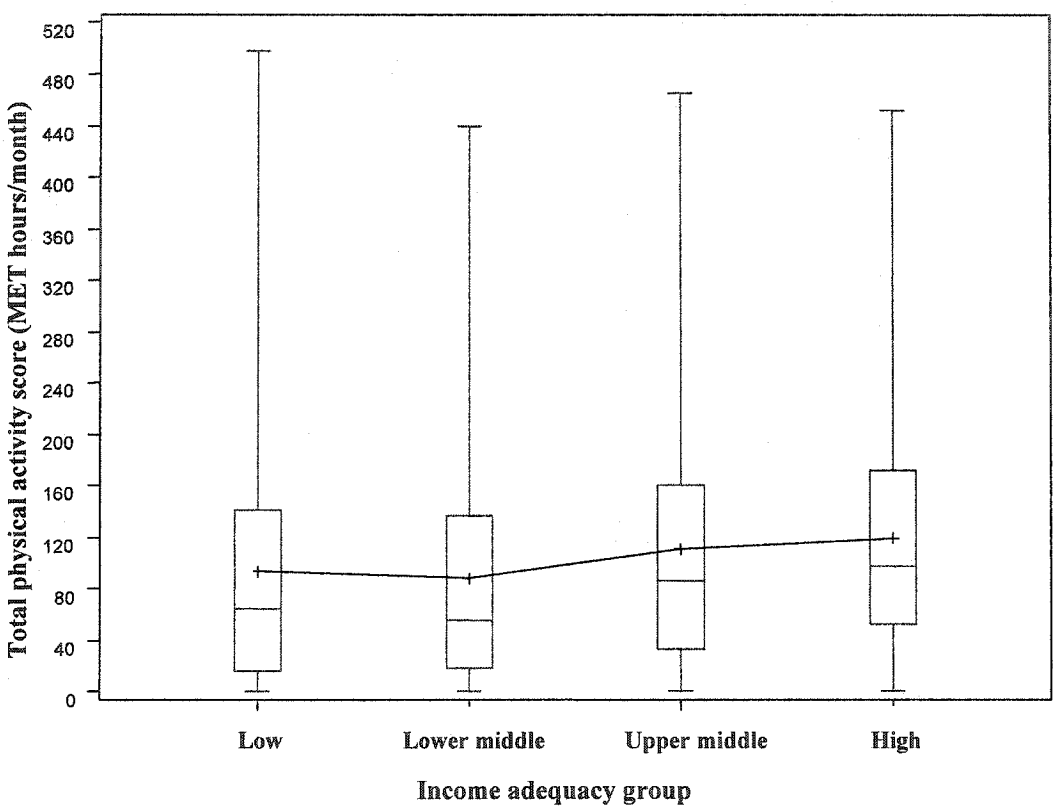
Box Plot B1-2 Total physical activity score by smoking group in male controls



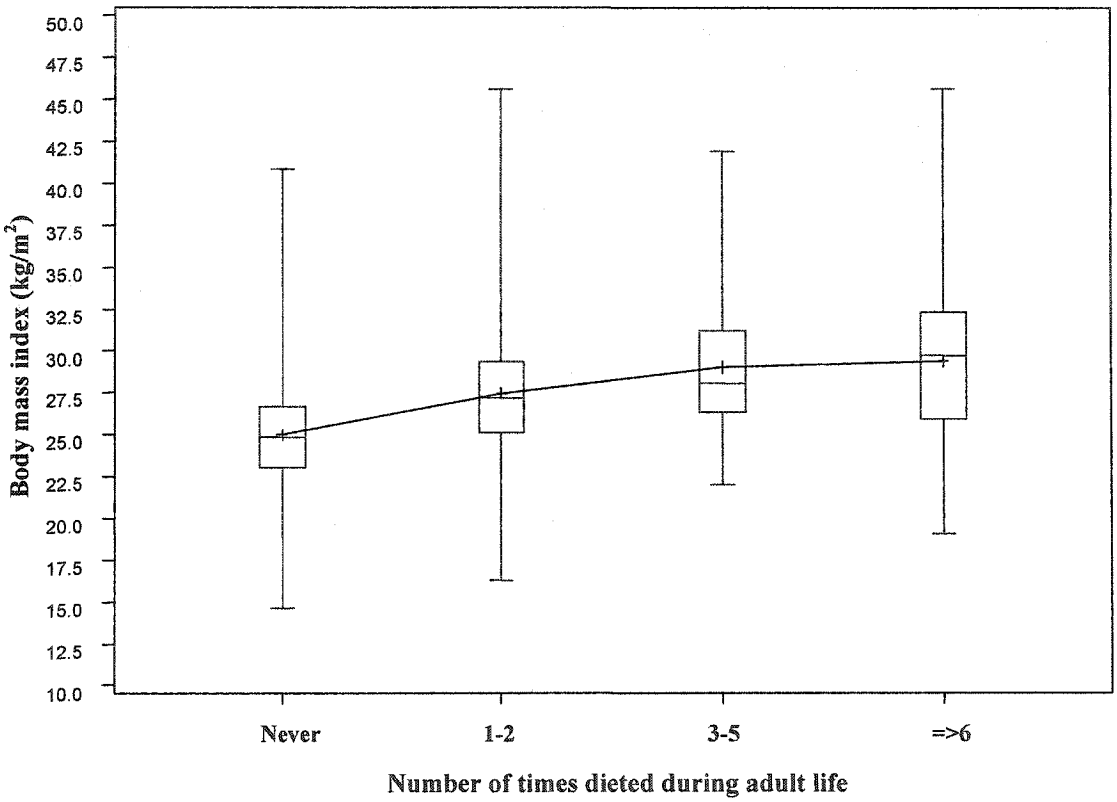
Box Plot B1-3 Total physical activity score by educational group in male controls



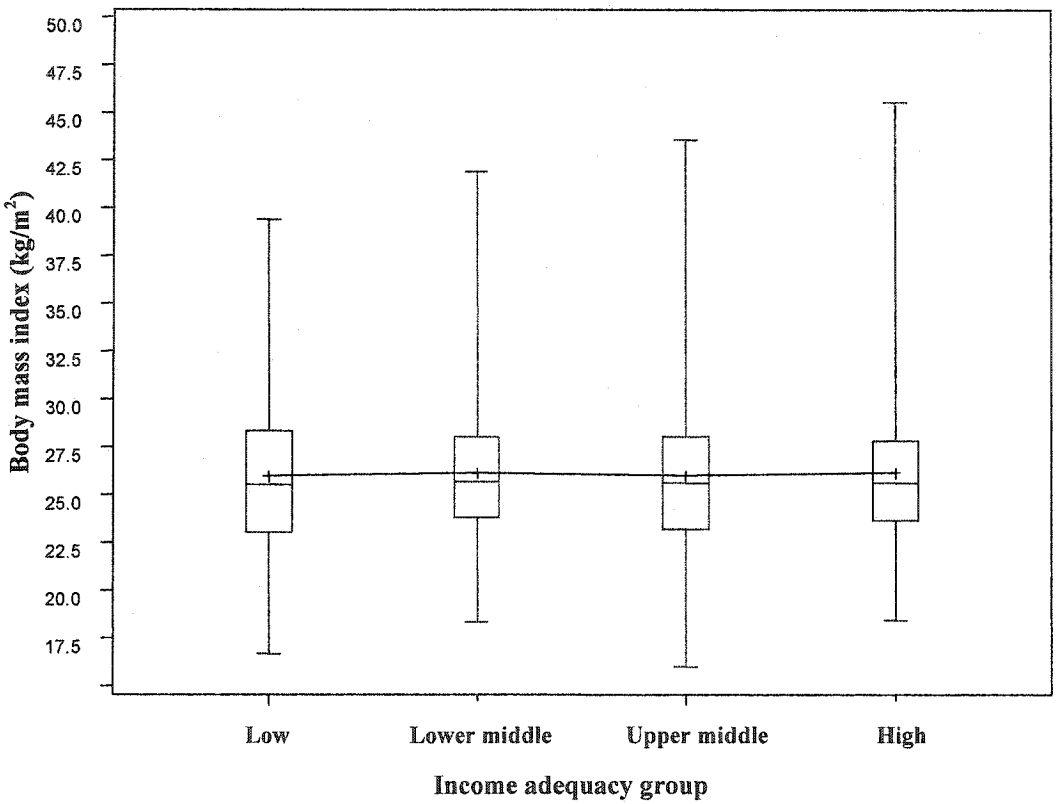
Box Plot B1- 4 Total physical activity score by income adequacy group in male controls



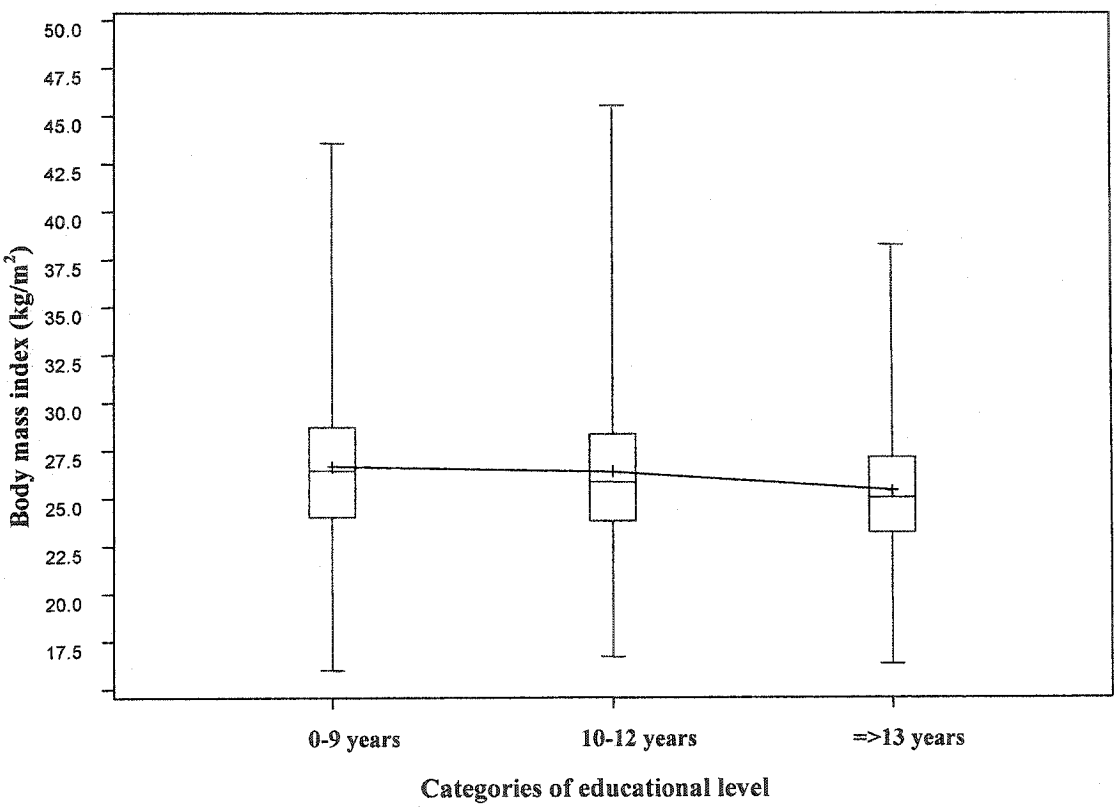
Box Plot B1-5 Body mass index by diet groups in male controls



Box Plot B1-6 Body mass index by income adequacy group in male controls

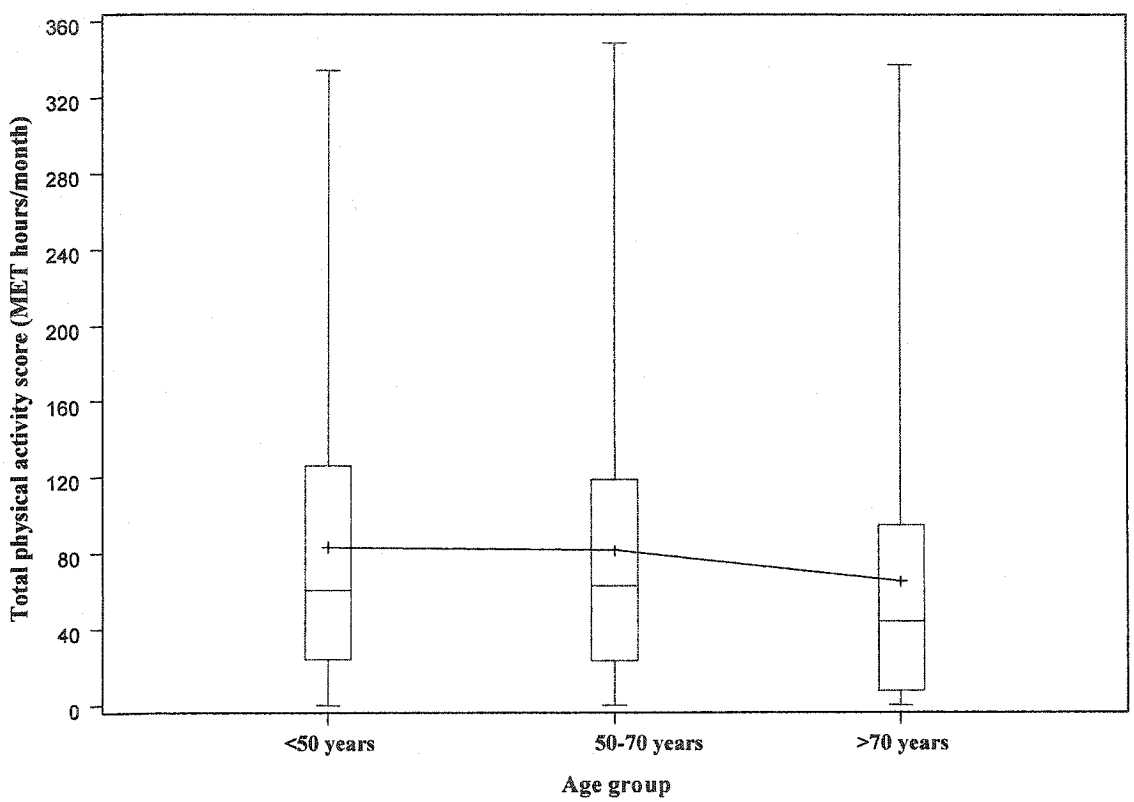


Box Plot B1-7 Body mass index by educational group in male controls

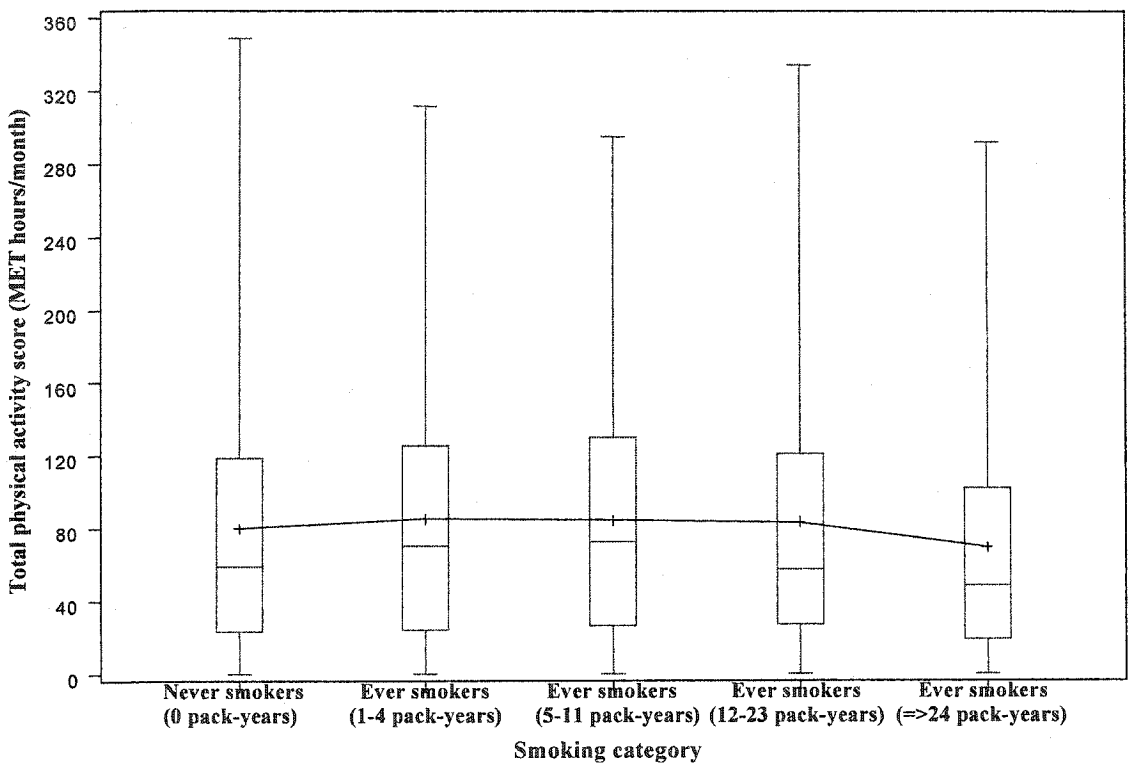


Appendix B2 – Female Controls

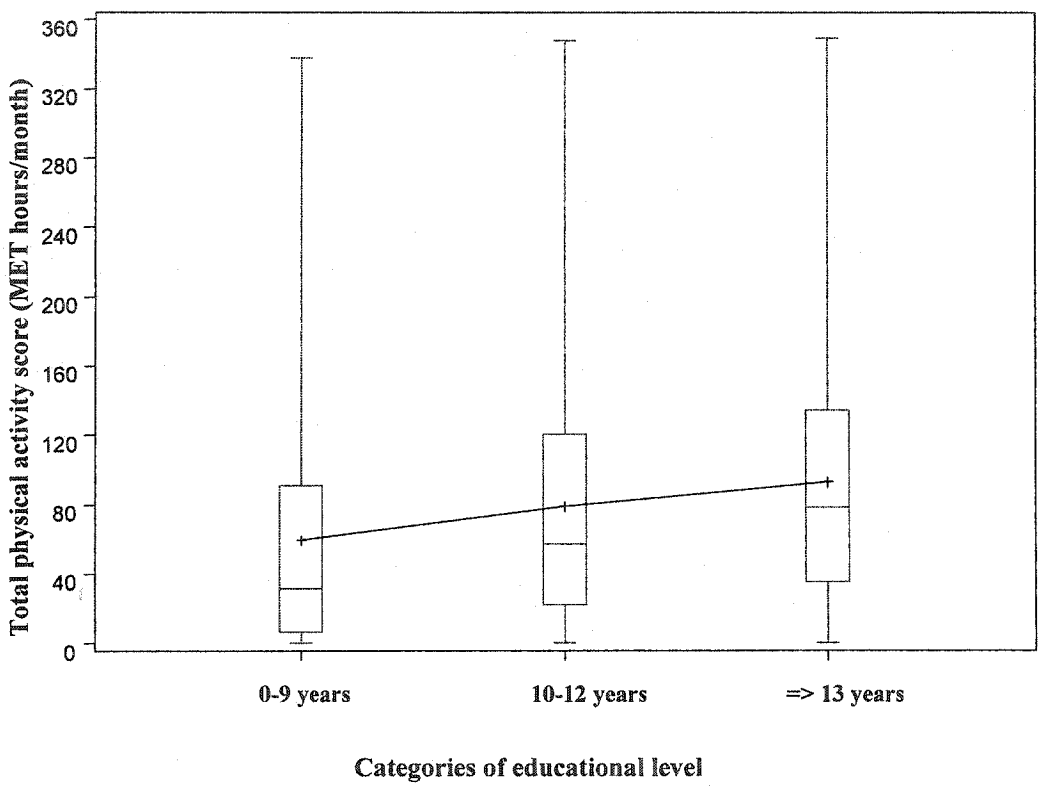
Box Plot B2 -1 Total physical activity score by age group in female controls



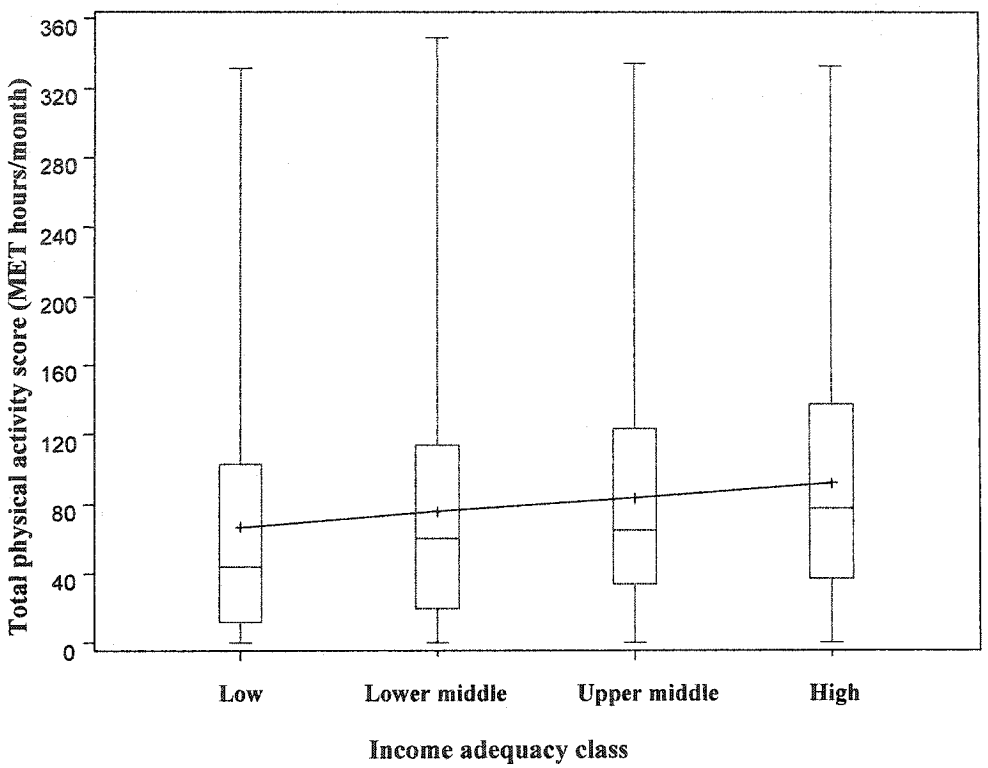
Box Plot B2-2 Total physical activity score by smoking group in female controls



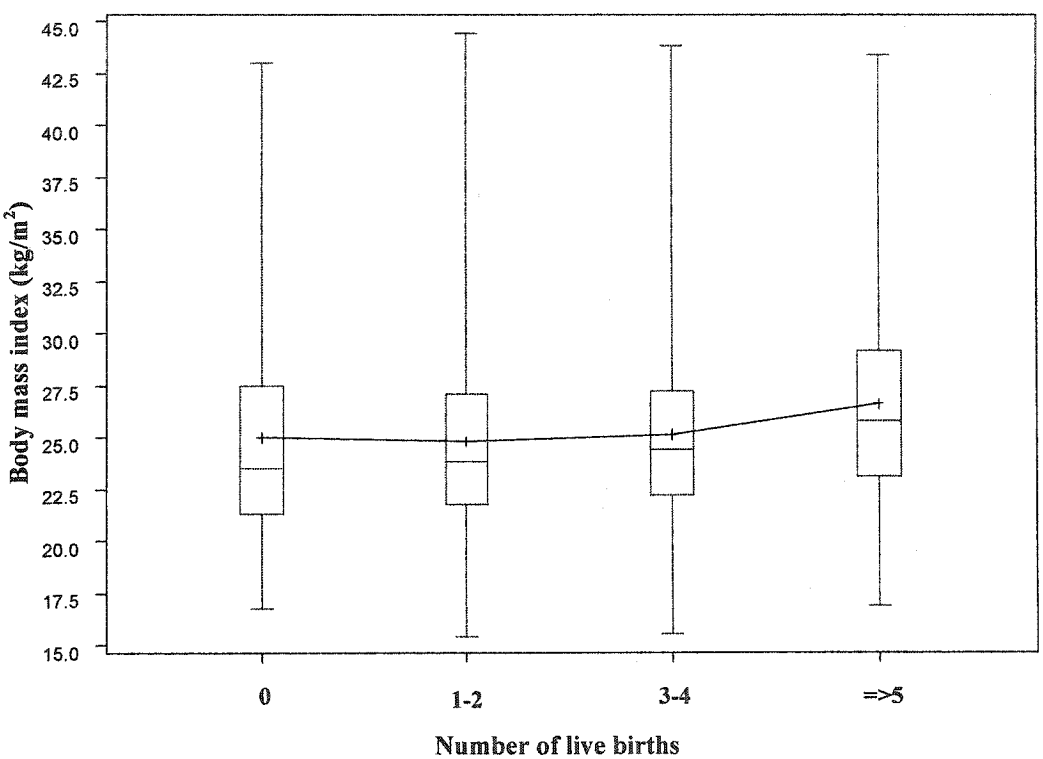
Box Plot B2-3 Total physical activity score by educational group in female controls



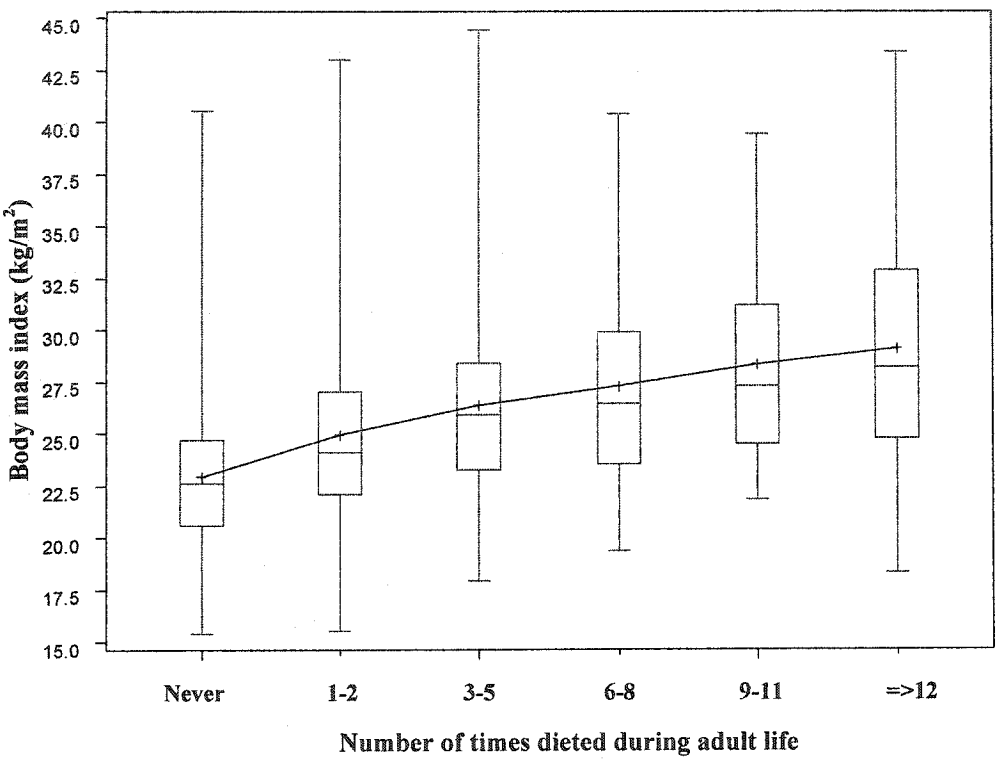
Box Plot B2-4 Total physical activity score by income adequacy class in female controls



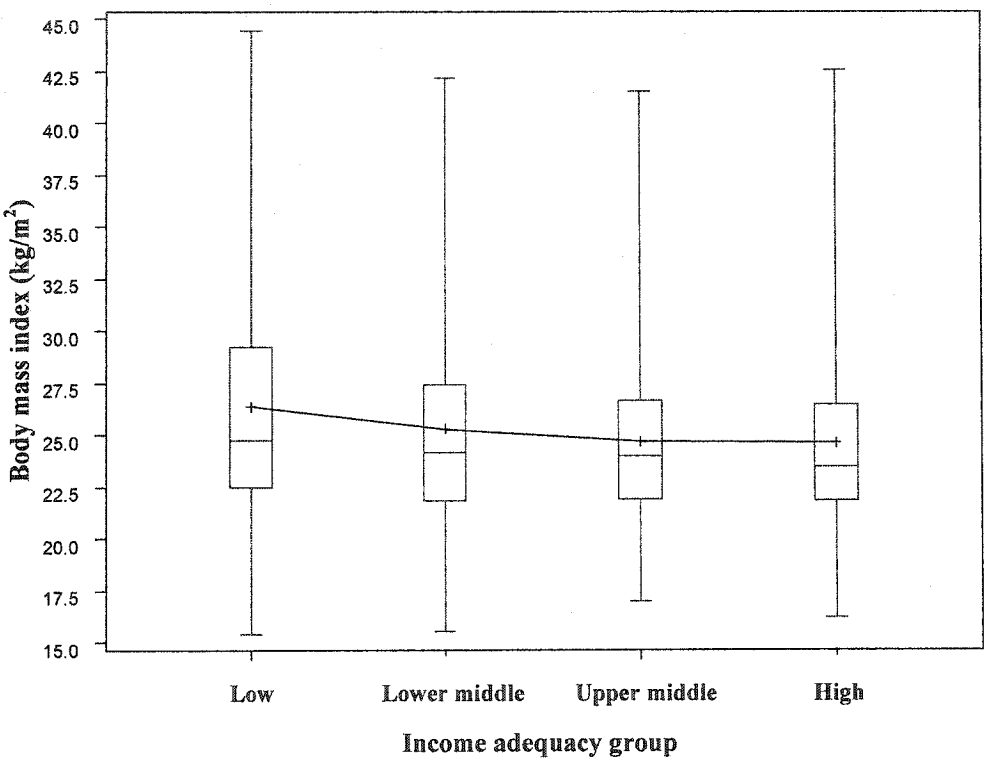
Box Plot B2-5 Body mass index by number of live births in female controls



Box Plot B2-6 Body mass index by diet group in female controls



Box Plot B2-7 Body mass index by income adequacy group in female controls



Box Plot B2-8 Body mass index by educational group in female controls