THE INFLUENCE OF ALCOHOL CONSUMPTION, SMOKING, AND PHYSICAL ACTIVITY ON PERIPHERAL BLOOD LEUKOCYTE TELOMERE LENGTH

by

Lidija Latifovic

A thesis submitted to the Graduate Program in Epidemiology in conformity with the requirements for the degree Master of Science

Queen’s University
Kingston, Ontario, Canada
June 2014

Copyright © Lidija Latifovic, 2014
Abstract

Background

Telomeres are repetitive DNA sequences and associated proteins that cap the ends of chromosomes. Functional telomeres protect genetic information and maintain chromosomal stability. Critically short telomeres are associated with an increased risk of cancer. Telomeres shorten at every DNA replication and due to the nature of their molecular structures are also particularly sensitive to oxidative damage. Consequently, oxidative stress accelerates age-dependent telomere shortening. Certain lifestyle factors affect the oxidant/anti-oxidant balance and can promote oxidative stress. The objectives of this thesis were to determine the effect of alcohol, smoking and physical activity on telomere length in a healthy adult population and to explore an interaction with CYP2E1 and alcohol on telomere length.

Methods

A cross-sectional study design was used to recruited 678 healthy volunteers from 2006 to 2008 from Kingston and Ottawa, Ontario, and Halifax, Nova Scotia. This sub-study included 477 individuals with an available DNA sample and complete questionnaire data. Alcohol consumption, smoking and physical activity were assessed by self-administered questionnaire and CYP2E1 was genotyped by the TaqMan® drug metabolism genotyping assay in the larger study. Relative leukocyte telomere length was measured using multiplex quantitative real-time PCR. Multiple linear regression was used to determine the effect of lifestyle on telomere length while controlling for important covariates. Effect modification by CYP2E1 genotype, specifically the CYP2E1*5B polymorphism, in relation to alcohol was investigated by the inclusion of a product term in the model.
Results

Alcohol consumption was not associated with telomere length. The interaction with CYP2E1 was not statistically significant; however, the effect of alcohol on telomere length was qualitatively stronger among those with the normal form of CYP2E1. Current smoking and pack-years smoking were inversely related to telomere length. After adjustment for confounders telomere length was not associated with total physical activity; however, higher vigorous physical activity was associated with longer telomeres.

Conclusions

In recent years, research has demonstrated a relationship between shorter telomeres and age-related diseases, including many forms of cancer. This study found that smoking was inversely associated and vigorous physical activity was positively associated with leukocyte telomere length.
Co-Authorship

This thesis is the work of Lidija Latifovic in collaboration with thesis supervisors Dr. Will D. King and Dr. Thomas E. Massey. The larger cross-sectional study was conceptualized, designed and executed by principal investigator Dr. Will D. King and co-investigator Dr. Thomas E. Massey, with funding from the Canadian Institutes of Health Research. Conceptualization of the thesis question was a collaborative effort between Lidija Latifovic and Dr. Will D. King. The measurement of telomere length in DNA isolated from peripheral blood leukocytes (with support from Sarah D. Peacock), statistical analysis, interpretation of results, and the writing of manuscripts were the work of Lidija Latifovic with supervision and contribution from Dr. Will D. King and Dr. Thomas E. Massey.
Acknowledgements

“Education sows not seeds in you, but makes your seeds grow.”

Khalil Gibran

Foremost, I would like to thank my thesis supervisors, Dr. Will King and Dr. Tom Massey, for their support, encouragement, and insightful guidance. Your complementary expertise made for stimulating discussions and an enriching graduate experience. Will, you have influenced my development as a researcher profoundly and I am truly grateful for having had the opportunity to work with you. Tom, thank you for always making my work stronger, without your consistent and illuminating instruction this thesis could not have reached its present form.

I would also like to thank the Queen’s Terry Fox Foundation Transdisciplinary Training Program in partnership with CIHR for promoting and financially supporting transdisciplinary science, enhancing my graduate training, and for continuously broadening my perspective on cancer research.

Sarah – may the PCR gremlins never bother us again! I thoroughly enjoyed our exchange of ideas and collaboration on the telomere part of this project and appreciated your company in the lab.

The faculty and staff in the Department of Public Health Sciences - thank you for the many ways you nurture an environment that encourages lifelong learning and great personal development.

Life would not have been the same without my classmates, particularly Shari Thomas and Erin Liu, who helped me stay fit and sane through both the bleakest and best parts of the past couple of years.

I would like to express my most heartfelt gratitude to my parents, Milica and Rasim, my sister Lejla, and my love A.J. Thank you for always loving, supporting, and encouraging me and for being a constant source of motivation. Thank you for always giving me the independence to pursue all that I set out to do in my own way. And finally, thank you for providing me with the security and sustenance that I need to ‘grow my seeds.’
Table of Contents

Abstract..............................................................................................................................................ii
Co-Authorship ......................................................................................................................................iv
Acknowledgements .......................................................................................................................... v
List of Figures ........................................................................................................................................i
List of Tables .........................................................................................................................................ix
List of Abbreviations ..........................................................................................................................xi
Chapter 1 Introduction ....................................................................................................................... 1
References........................................................................................................................................... 6
Chapter 2 Literature Review ................................................................................................................. 9
  2.1 Introduction................................................................................................................................. 9
  2.2 Telomere Structure and Function ............................................................................................ 9
  2.3 Telomere Length in Ageing and Cancer ..................................................................................10
  2.4 The Epidemiology of Telomere Length and Cancer .................................................................13
  2.5 Methods for Measuring Telomere Length .............................................................................14
  2.6 Surrogate Tissue Telomere Length .........................................................................................16
  2.7 Telomere Length Variation .......................................................................................................16
  2.8 Factors Affecting Telomere Length .........................................................................................17
  2.9 Oxidative Stress and Inflammation .........................................................................................18
  2.10 Alcohol Metabolism, Cigarette Smoking, Physical Activity and Oxidative Stress ..........19
      2.10.1 Alcohol Metabolism ....................................................................................................... 19
      2.10.2 Cytochrome P450 2E1 (CYP2E1) ...................................................................................20
      2.10.3 Cigarette Smoking ........................................................................................................21
      2.10.4 Physical Activity ............................................................................................................22
  2.11 The Epidemiology of Alcohol Consumption, Cigarette Smoking, Physical Activity and Telomere Length ................................................................................................................23
      2.11.1 Alcohol Consumption ...................................................................................................23
      2.11.2 Cigarette Smoking ..........................................................................................................24
      2.11.3 Physical Activity ............................................................................................................24
  2.12 Minimum Detectable Effects ....................................................................................................25
  2.13 Ethical Consideration ..............................................................................................................26
  2.14 Study Objectives and Hypothesis ...........................................................................................27
References........................................................................................................................................... 30
List of Figures

Figure 2-1. Mechanism of telomere shortening due to the end-replication problem. .................. 28
Figure 2-2. A conceptual diagram of the proposed mechanisms relating selected lifestyle and
 genetic factors to telomere length and cancer. ........................................................................... 29
Figure 3-1. Multiplex qPCR assay specificity verification by melt curve analysis showing
telomere and betaglobin products. .................................................................................................... 61
Figure 3-2. Gel showing the single copy gene (hbg) and telomere (tel) products. .................. 62
Figure 3-3. Example of standard curves for betaglobin and telomere reactions. ...................... 63
Figure 5-1. Alcohol metabolism via the ADH-ALDH pathway. ................................................... 98
Figure 5-2. CYP2E1 metabolism of alcohol. ................................................................................. 99
List of Tables

Table 3-1. Distribution of samples with no amplification or weak qPCR signal by age category. 64
Table 3-2. Reproducibility of telomere length measurement calculated using one-way analysis of variance. ........................................................................................................................................65
Table 4-1. Characteristics of the study population and age-adjusted rLTL. 82
Table 4-2. Results from multivariable linear regression presenting adjusted regression coefficients relating lifestyle exposure variables to rLTL. ............................................................................................................. 83
Table 5-1. Age-adjusted regression coefficients and associated standard errors. 100
Table 5-2. Exploratory analysis of effect modification by the CYP2E1*5B polymorphism on the relationship between alcohol consumption and rLTL. .................................................................................................................. 101
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’</td>
<td>Five prime</td>
</tr>
<tr>
<td>$X^2$</td>
<td>Chi-square</td>
</tr>
<tr>
<td>µl</td>
<td>Microliter</td>
</tr>
<tr>
<td>A</td>
<td>Adenosine</td>
</tr>
<tr>
<td>ADH</td>
<td>Alcohol dehydrogenase</td>
</tr>
<tr>
<td>ALDH</td>
<td>Aldehyde dehydrogenase</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CIHR</td>
<td>Canadian Institutes of Health Research</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of variation</td>
</tr>
<tr>
<td>CYP2E1</td>
<td>Cytochrome P450, family 2, subfamily E, polypeptide 1</td>
</tr>
<tr>
<td>CYP2E1*5B</td>
<td>rs3813867, a PstI/RsaI polymorphism, caused by a single nucleotide G&gt;C change at position 1293 in the 5’-flanking region of the CYP2E1 gene; the G and C alleles are also known as c1 and c2 respectively</td>
</tr>
<tr>
<td>DC</td>
<td>Dyskeratosis congenita</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence in situ hybridization</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>Hbg</td>
<td>Human betaglobin</td>
</tr>
<tr>
<td>HDL</td>
<td>High-density lipoprotein</td>
</tr>
<tr>
<td>IPAQ</td>
<td>International physical activity questionnaire</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium chloride</td>
</tr>
<tr>
<td>LDL</td>
<td>Low-density lipoprotein</td>
</tr>
<tr>
<td>rLTL</td>
<td>Relative leukocyte telomere length</td>
</tr>
<tr>
<td>MET</td>
<td>Metabolic equivalent of task</td>
</tr>
<tr>
<td>MIQE</td>
<td>Minimum information for publication of quantitative real-time PCR experiments</td>
</tr>
<tr>
<td>NS</td>
<td>Nova Scotia</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ODM</td>
<td>Ontario Drug Monitor</td>
</tr>
<tr>
<td>ON</td>
<td>Ontario</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PCR-RFLP</td>
<td>Polymerase chain reaction-restriction fragment length polymorphism</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>SAS</td>
<td>Statistical Analysis System</td>
</tr>
<tr>
<td>Scg</td>
<td>Single-copy gene</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SE</td>
<td>Standard error</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SST</td>
<td>Serum-separating tube</td>
</tr>
<tr>
<td>STELA</td>
<td>Single telomere length analysis</td>
</tr>
<tr>
<td>SYBR Green I</td>
<td>Registered trademark of Synergy Brands, Inc.; cyanine dye that binds double-stranded DNA to detect PCR product as it accumulates during the reaction</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
<tr>
<td>Tel</td>
<td>Telomere</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane, formula (HOCH₂)₃CNH₂</td>
</tr>
<tr>
<td>T/S ratio</td>
<td>Telomere to single-copy gene ratio</td>
</tr>
<tr>
<td>UV-Vis</td>
<td>Ultraviolet-visible</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

According to the Canadian Cancer Statistics (1) the lifetime probability of developing cancer in Canada is 40% and the lifetime probability of dying from cancer is 28%. Cancer is the leading cause of death in Canada, causing 29.8% of all deaths. In addition to the personal ramifications of cancer it has a major economic impact on society costing $17.4 billion in both direct and indirect costs (1). As Canada’s population ages the burden of cancer will continue to increase; building on current success in prevention and cancer control remains crucial. According to the World Cancer Research Fund (2), 30% of cancers are preventable through lifestyle modifications such as maintaining a healthy diet and healthy weight, reduced alcohol consumption, smoking cessation, and regular physical activity.

Risk of cancer is determined by a complex interplay between environmental exposures and genetic and nutritional factors that regulate responses to carcinogens (3). One approach of molecular epidemiology in cancer control is the utilization of biomarkers to either identify harmful exposures or to identify populations at risk for cancer in time to effectively intervene. This includes biomarkers of exposure that better approximate the effective dose of an exposure, biomarkers of susceptibility, factors that may make certain individuals more sensitive to an exposure, or biomarkers of early effect, indicators of early sub-clinical changes. Thus, biomarkers may be valuable to both primary and secondary cancer prevention.

A further advantage of biomarkers of early effect in research is that they allow the study of exposure-outcome relationships in healthy populations, the study of an outcome that occurs more frequently and on a shorter temporal scale than the disease itself, which is particularly
advantageous when studying diseases of long latency such as cancer, and the elucidation of the mechanism of action relating the exposure to the disease (3).

While one might argue that from a public health point of view understanding the mechanism of lifestyle-mediated carcinogenesis is less important than the knowledge that lifestyle factors can cause cancer, it can also be argued that many people will not alter their behaviour even if they are aware of the associated increased risk of disease. Understanding the mechanism of carcinogenesis allows for the identification of early stages in the disease process when it is still possible to prevent the full onset of disease (4). Well-validated biomarkers allow for the identification of individuals at high risk and the potential to interrupt the natural progression of the disease before it has fully developed.

Telomere length is one such potential tool in cancer prevention. Telomeres are repeating DNA-protein complexes that function to maintain genetic integrity; however, telomeres shorten as cells divide exhibiting a progressive age-dependent attrition. Short telomeres trigger genetic instability and telomere dysfunction is an early event in cancer progression (5). At least a few hundred nucleotides of telomeric DNA and associated proteins must cap the ends of chromosomes to maintain function (6, 7). When too many short telomeres accumulate the cell either executes programmed cell death or enters senescence – a state where further cell division is no longer possible. This process contributes to a loss of tissue and organ function with age (8). Therefore, short telomeres may promote carcinogenesis both by triggering genetic instability and by contributing to senescence, which can affect tissue function. The most consistent evidence for an association with short telomeres exists for lung, bladder, esophageal, gastric, ovarian, head and neck, and renal cancers (9). The published evidence is suggestive but not conclusive for Non–Hodgkin lymphoma, breast, and colorectal cancer (9).
Lifestyle factors, including alcohol consumption, smoking, and physical activity, are modifiable and prevalent exposures and their importance in the etiology of cancer has been established (10). The International Agency for Research on Cancer (IARC) has classified alcoholic beverages as carcinogenic to humans and has declared a causal relation between the consumption of alcoholic beverages and cancers of the oral cavity, pharynx, larynx, oesophagus, liver, colorectum and breast (12). A meta-analysis of 156 studies demonstrated a clear dose-response relationship between daily alcohol consumption and increased risk of cancer (13). Despite this, the majority of Canadians are unaware of the link between alcohol use and cancer and at least 9.2% of Canadians exceed published low-risk drinking guidelines (11). The relationship between alcohol and cancer is further complicated by polymorphic differences in genes involved in alcohol metabolism. Genetic variation that leads to a functional change in the encoded protein can create differences between individuals with respect to the carcinogenic exposure resulting from alcohol intake. Consideration of an interaction between alcohol intakes and common genetic polymorphisms involved in alcohol metabolism may inform on the alcohol-cancer relationship.

Approximately 15% of all incident cases of cancer in Ontario are attributable to tobacco (14). Actively smoking cigarettes causes cancer in over twenty different tissue types. Specific population attributable fractions for smoking in Canada include 71% of lung and bronchus, 36% of bladder, 41% of esophageal, 26% of liver, 16.3% of kidney, 19.4% of stomach, and 11% of colorectal cancers (14). The age-standardized prevalence of current smoking in Canada is 16.7%, equating to roughly 4.7 million people (14). Despite an expected long-term decrease in smoking rates, absolute numbers of current smoking are expected to increase in future years due to ongoing population growth (14). Cancer prevention strategies aimed at reducing the use of
tobacco have been ongoing since 1964 (16). Nevertheless, some 3 million deaths a year are estimated to be attributable to smoking, an estimate that is expected to increase to 10 million a year in 30-40 years (17, 18).

Physical activity may contribute to cancer prevention; convincing evidence exists for a decreased risk with increased physical activity for colon and pre- and postmenopausal breast cancers, probable for prostate cancer, possible for lung and endometrial cancers and insufficient for cancers at other sites (19). According to data from the Canadian Health Measures Survey (22), 15% of adults, 17% of men and 14% of women, accumulate 150 minutes of moderate to vigorous physical activity per week in 10 minutes bouts, as per Canadian Physical Activity Guidelines (20, 21); however, only 5% do so on a regular basis of at least 30 minutes on at least 5 days per week.

The objectives of this thesis were to determine the effect of alcohol consumption, an interaction between alcohol and the CYP2E1*5B variant, smoking, and physical activity on telomere length, a biomarker of increased cancer risk, using a nested cross-sectional design. The larger study collected questionnaire data and a blood sample for biomarker analysis for 678 participants; however, this thesis included 477 participants with available DNA for whom leukocyte telomere length was measured using multiplex qPCR. Alcohol consumption and cigarette smoking are hypothesized to accelerate age-related telomere shortening and are expected to be associated with shorter telomere length, while physical activity is hypothesized to attenuate oxidative stress and is expected to be associated with longer telomere length. Compared to the normal CYP2E1*1A, the CYP2E1*5B variant may metabolize alcohol faster and generate more reactive oxygen species making a greater contribution to the generation of oxidative stress (23). Thus, the effect of alcohol on telomere length is expected to be greater among those with the CYP2E1*5B variant.
This document conforms to the guidelines for a manuscript-based thesis as recommended by the School of Graduate Studies at Queen’s University. The second chapter reviews current literature on telomere structure and function, the role of telomere length in cancer initiation and progression, and the contribution of alcohol, CYP2E1, smoking and physical activity to oxidative stress. The third chapter describes in detail the development, applications and evaluation of telomere measurement in the study. This represents a substantial and unique contribution by the study investigator and is not covered in detail in the manuscript. The fourth chapter presents the main results of this study in the form of a manuscript, which has been prepared for submission in accordance with the guidelines set out by Cancer Epidemiology, The International Journal of Cancer Epidemiology, Detection and Prevention. Chapter five presents the exploratory investigation of the potential effect modification by CYP2E1 genotype on the relationship between alcohol and telomere length. Finally, chapter six summarizes the main findings, considers strengths and limitations of the research and discusses implications and future directions.
References


Chapter 2

Literature Review

2.1 Introduction

Telomeres are repetitive DNA sequences and associated proteins located at the ends of linear chromosomes. They serve as “caps” that protect the ends of chromosomes from degradation, fusion, and breakage, and serve to prevent the loss of genetic material and maintain genetic stability. As cells proliferate telomeres shorten in an age-dependent manner, some reaching critically short lengths that render them dysfunctional. Short telomeres are associated with numerous age-related diseases including cancer at several sites. The most consistent evidence exists for lung, bladder, esophageal, kidney, head and neck, ovarian, and stomach cancer. A growing body of literature demonstrates that the environment and lifestyle can affect the rate of telomere shortening (1). Alcohol consumption, smoking and physical activity are modifiable aspects of lifestyle that are risk factors for cancer at numerous sites. Lifestyle factors may act in different ways to promote cancer development; one potential mechanism is through oxidative stress and inflammation, which can subsequently impact telomere length. This review describes telomere structure and function, the role of short telomeres in cancer initiation and progression, the contribution of oxidative stress and inflammation to telomere attrition, and the influence of alcohol, CYP2E1, smoking and physical activity on oxidative stress.

2.2 Telomere Structure and Function

Telomeres are repeating TTAGGG DNA sequences located at both ends of human linear chromosomes. They bind shelterin complexes, six telomere associate proteins that protect chromosome ends from degradation and non-homologous recombination (2, 3). Human telomeres
end in a 150-200 nucleotide single-strand overhang which can fold back to form a large telomeric loop known as the T loop (4). Formation of the T loop has been proposed as the mechanism by which telomeres protect chromosome ends from recognition as double-strand breaks in DNA. Damaged DNA normally sets off a cellular response that stops cell growth to allow for DNA repair by end-to-end joining (5). Inappropriate joining of chromosome ends creates unstable chromosomes that break during cell division resulting in genomic rearrangements. Functional telomeres protect the cell from the degenerative processes associated with inappropriate DNA repair.

In addition to protecting chromosomes from inappropriate DNA repair, telomeres prevent chromosome degradation during DNA replication. DNA replicative machinery is unable to completely copy the ends of linear chromosomes, a phenomenon known as the end-replication problem (6). Without telomeres this would lead to a loss of genetic information at every cell division. During early human development and in germ line cells, the enzyme telomerase elongates and maintains telomere length by adding TTAGGG repeats to telomere ends (7). However, telomerase activity is insufficient in most adult human somatic cells (8,9) and telomere DNA that is lost at each cell division is not replaced, resulting in a progressive, age-dependent shortening of telomeres (10). In normal cells, progressive telomere erosion results in dysfunctional telomeres that no longer bind telomere-capping proteins, signaling the activation of a persistent DNA damage response that initiates either senescence - an end to cell replication, or programmed cell death (11).

2.3 Telomere Length in Ageing and Cancer

Since first proposed by Harley et al. (6) a great deal of evidence has emerged implicating telomere shortening as a determinant in human disease. Many studies have shown an inverse
association between telomere length and age; likewise, short telomeres are characteristic of several age-related diseases (12-15). Certain hereditary diseases associated with mutations in telomere maintenance suggest a mechanistic role of telomere length in ageing. Individuals with premature ageing syndromes such as dyskeratosis congenita, aplastic anemia, and Werner’s syndrome have short telomeres, increased chromosomal instability and accelerated cellular senescence (16). Most healthy human tissues also show significant telomere shortening with increasing age (17, 18). The strength of the association with age is highly dependent on the age range of the population under study (19); however, leukocyte telomere length is consistently associated with age across populations, measurement methods and statistical models (summarized in 19).

A strong link exists between advanced age and increased incidence of cancer. Several age-related molecular and physiological changes, including mutation load, epigenetic regulation and modification of the cellular microenvironment, may act synergistically with telomere dysfunction to promote cancer (20). In response to short telomeres cells enter senescence, halting cell cycle progression, preventing further division, and entering a state of irreversible growth arrest. This process can function to suppress cancer by preventing the proliferation of damaged cells and the resulting accumulation of somatic mutations that promote carcinogenesis. However, senescence also has other complex and opposing effects including the promotion of ageing and tumour progression (21). There is mounting, although not yet definitive evidence, that suggests senescence may contribute to ageing by two main mechanisms: an accumulation of senescent cells in tissues, to a point that compromises functionality, and by limiting the regenerative potential of stem cells resulting in a loss of stem cell repair and replenishment of worn out and damaged tissue (22). Consequently, the accumulation of senescence may promote functional
decline in many tissues contributing to the onset of age-associated degenerative diseases including cancer.

Senescent cells no longer divide but can remain viable and metabolically active for many years (23). Senescence-associated secretory pathways secrete factors such as metalloproteases, growth factors and potent inflammatory cytokines that disrupt tissue structure and function and promote chronic inflammation (21). While senescence protects against cancer earlier in life, the accumulation of senescent cells with age may create a microenvironment that supports malignant transformation (24). Additionally, the increased cell turnover induced by chronic inflammation may contribute to telomere shortening (24).

In the presence of certain genetic and epigenetic changes cells may ignore the DNA damage signal generated by critically short telomeres, evade either senescence or cell death, and continue to divide eventually entering crisis, characterized by genetic instability and extensive chromosomal abnormalities (24, 25). The vast majority of cells in crisis die; however rarely a few cells escape crisis by reactivating telomerase, which stabilizes telomeres, enabling continued cell division and growth (4, 23, 26). Reactivation of telomerase is detected in 90% of human cancers (27).

Telomere length dysfunction occurs in most human epithelial cancers as precursor lesions in many tissues display shorter telomeres than adjacent normal tissue (28). For example, Roger et al. (29) demonstrated that initiation of colon cancer is preceded by widespread telomere erosion and that in combination with other genetic changes telomere shortening caused large-scale genomic instability in polyps from familial adenomatous polyposis patients. Studies in breast cancer found that telomere erosion progressed through the different stages of breast cancer.
Development and transition from usual ductal hyperplasia to carcinoma *in situ* was marked by an increase in chromosomal abnormalities and telomerase reactivation (30).

### 2.4 The Epidemiology of Telomere Length and Cancer

Telomere length in non-cancer tissues, such as blood or buccal cells, has been associated with cancer and other diseases. Early evidence suggesting a relationship between telomeres and cancer derives from studies of patients with dyskeratosis congenita (DC) and direct evidence is derived from two meta-analyses summarizing the relationship with telomere length at different cancer sites. Patients with DC, a rare inherited syndrome marked by defects in telomere maintenance, have very short telomeres for their age, defined as below the first percentile among 400 healthy control subjects from birth to 100 years of age (31). The cumulative incidence of cancer among patients with DC is 40% by age 50 and 60% by age 68. DC patients have an eleven-fold increased risk of overall cancer compared to the general population with a median age at cancer diagnosis of 29 years (32). Although DC represents extremely short telomere lengths, it is likely that less pronounced differences in telomere length in the healthy population also contribute to cancer risk.

Cancer risk in relation to age-adjusted surrogate tissue telomere length, mainly blood and buccal cells, has been investigated in 27 studies and for 13 cancer sites, which are summarized in two recent meta-analyses (33, 34). Both of these meta-analysis concluded that shorter telomere length is associated with increased overall incident cancer risk and reported a summary OR of 1.96 (95% CI 1.37-2.81) (33) and 1.35 (95% CI 1.14-1.60) (34) for cancer risk in those with shorter telomeres versus longer telomeres.

The most consistent evidence for a relationship with shorter telomere length exists for bladder cancer (OR range for shortest compared to longest telomere length =1.88 – 4.5; based on
3 studies), esophageal cancer (OR=2.52 – 4.66; 2 studies), gastric cancer (OR=2.04 – 3.12; 2 studies), head and neck cancer (OR=5.11, 95% CI: 1.90 - 13.77; 1 study), ovarian cancer (OR=4.89, 95% 1.93 – 12.34; 1 study), and renal cancer (OR=4.41 – 5.26; 2 studies). The evidence for a relationship with Non-Hodgkin lymphoma, breast cancer, and colorectal cancer is inconsistent. Based on results from single studies short telomere length does not appear to be associated with endometrial, prostate and skin cancers (33, 34).

In summary, a strong association has been demonstrated at several cancer sites, consistently in different populations and using different methods. Results from functional studies provide a convincing biological mechanism for the role of short telomeres in cancer initiation and progression and both animal and epidemiologic studies have shown that shorter telomeres are associated with an increased risk of cancer (35-38). However, most of the studies published on the relationship are retrospective case-control studies where DNA samples from the cases were collected after cancer diagnosis, rendering the studies vulnerable to reverse causation bias. One prospective study (39) on the relationship concluded that age-adjusted telomere length was predictive of increased cancer risk (OR: 2.15 95% CI: 1.12-4.14 for middle telomere length and OR: 3.11 95%CI: 1.65-5.84 for the shortest telomere length compared to the longest telomere length group).

2.5 Methods for Measuring Telomere Length

In epidemiological studies telomere length is most often measured in leukocyte DNA derived from peripheral blood using the quantitative polymerase chain reaction (qPCR). However, several other methods exist for measuring telomere length. Repeatability and reproducibility are method dependent and inter-assay variation remains an issue; although, this variation may be reduced with good quality control (176). The Southern blot technique for
measuring telomere restriction fragment (TRF) lengths is currently considered the ‘gold standard.’ This method requires a greater amount of DNA and includes the non-canonical portion of telomeres which can vary in individuals and influence the obtained telomere length measure. Nevertheless, when performed proficiently Southern blot generally has good precision and inter-assay variability is reported to be >2% (44). Single telomere length analysis (STELA) and quantitative fluorescence in situ hybridization (qFISH), measure only the canonical region of telomeres and can also measure the spectrum of telomere lengths on each chromosome, the telomere length in individual cells and for different cell types. Thus, STELA and qFISH can quantify the proportion of short telomeres in a sample (40). Inter-assay variability for qFISH and STELA is generally >5% (177).

The advantages of the qPCR method for measuring telomere length are that it is fast, sensitive, less resource intensive, and requires much less DNA making it easily scalable for high-throughput analysis (41). Inter-assay variation for qPCR, which also only measures the canonical region of telomeres, is generally >6% (44). While STELA is usually restricted to several well-characterized chromosomes and qFISH requires viable, metaphase-arrested cells, qPCR can measure telomere length in genomic DNA isolated from blood cells, a tissue that is relatively easily obtained with minimal risk to participants.

The qPCR method for measuring telomere length was developed by Cawthon and recently updated to multiplex qPCR (42). It measures telomere repeats relative to a control single copy gene resulting in an estimate of telomere length in the sample cell population that is expressed in T/S ratio units. The T/S ratio for each sample is standardized to a reference DNA sample to allow comparability between reactions. The validity of the multiplex qPCR has been demonstrated by Cawthon (42) and others (43, 44) in comparison to the Southern blot method for
measuring telomeres. Relative telomere length measured by qPCR is strongly correlated to TRF lengths from Southern blot with correlation coefficients ranging from 0.85 to 0.92 (42-44). Additionally, the short-term stability of telomere length was established by Kim et al. (45), who reported an interclass correlation coefficient of 0.64 for telomere length measured in blood samples collected at seven visits spanning a nine-month period.

2.6 Surrogate Tissue Telomere Length

Studies comparing telomere length in different human tissues have reported that while telomere length differs between tissue types it is highly correlated within the individual. Telomere length is highly synchronized in fetal tissue (46) and at birth among leukocytes, umbilical artery cells and skin cells (47). Results from studies in adult populations corroborate those in newborns; adult leukocyte telomere length is correlated with telomere length in fibroblasts (48), skin (49-51), synovial tissue (50), skeletal muscle and subcutaneous fat (51). Leukocyte telomere length represents the average telomere length across leukocytes, a heterogeneous population of cells, which have different average telomere lengths (52). However, there is robust synchrony among leukocyte subsets throughout the lifespan (53) and inter-individual differences are much greater than intra-individual differences in leukocyte telomere length (54).

2.7 Telomere Length Variation

Leukocyte telomere length ranges from 10 – 18 kb at birth (55) and approximately 50-100 bp are lost at each cell division (56). Telomere shortening in peripheral blood leukocytes varies with age, with accelerated shortening in childhood and old age and a linear loss in adolescence and adulthood (55). There is considerably more inter-individual variation in telomere length measured in adult leukocytes than in neonates (55) suggesting a considerable impact of the environment on telomere length.
Under a state of high oxidative stress, an accumulation of single-strand breaks in DNA is the major cause of telomere shortening (57, 58). DNA damage induced by oxidative stress has been extensively studied and described as a major mechanism involved in accelerated telomere shortening (57-61, 63). Due to the high content of guanine (G) nucleotides in their repeating sequences, telomeres are particularly sensitive to the formation of oxidative bases (64, 65). Single-strand breaks caused either directly by reactive oxygen species or indirectly as part of the DNA repair process are not efficiently repaired in telomere DNA (62, 66) and *in vitro* studies demonstrate an up to ten times increased rate of telomere attrition under oxidative stress (60, 61), which is attenuated by treatment with antioxidants (58, 67).

Inflammation can promote oxidative stress; further, inflammation increases the rate of blood stem cell replication because of the increased demand for leukocytes in the inflammatory process (68). A chronic increase in the systemic burden of oxidative stress and inflammation enhances the rate of telomere shortening in blood stem cells, which is ultimately expressed in shortened leukocyte telomere length (54). Therefore, oxidative stress may accelerate telomere shortening by damaging DNA. Inflammation may accelerate telomere shortening indirectly by promoting oxidative stress and by increasing the demand for leukocyte cell replication. Evidence indicates that oxidative damage accumulates with age (69) and oxidative stress is a major contributor to age-related diseases such as cardiovascular disease (87), pulmonary diseases (88), diabetes (89), neurodegenerative diseases (90) and cancer (91).

### 2.8 Factors Affecting Telomere Length

Currently determinants of telomere length are poorly understood. While leukocyte telomere length is to some extent heritable and determined by genetics (70-74) it is also determined by environmental and lifestyle factors (1). Older age, male gender and white race
were consistently associated with shorter leukocyte telomere length across different populations, measurement methods and statistical models (reviewed in 19). Uncertainty still exists with other risk factors; however, limited evidence suggests that unsaturated fatty acid intake, particularly linoleic fatty acid (75), smoking, and obesity (75, 76) are inversely associated with telomere length. While higher vitamin intake and multivitamin use (77), vitamin D biomarkers (78), cereal fiber intake (75), and physical activity (79) are associated with longer telomeres. Less education (80), lower socioeconomic status (81), greater life stress (82), and type II diabetes (83) have also been associated with shorter telomeres. These studies have proposed oxidative stress and inflammation as the mechanisms mediating these relationships.

2.9 Oxidative Stress and Inflammation

Under normal physiological conditions the body produces potentially dangerous oxidants as a byproduct of normal aerobic metabolism. Reactive oxygen species and reactive nitrogen species (ROS and RNS), referred to as ROS here for simplicity, play a dual role in biologic systems and can be both beneficial and harmful. ROS function in cellular signaling and immune response; however, at high concentrations ROS cause damage to lipid membranes, proteins and DNA (84). Illness, inflammation, lifestyle factors, exposure to radiation, ischemia/reperfusion and various enzymes involved in cell functioning leak electrons and generate ROS (85). The body relies on antioxidant systems to control the effects of ROS; however, when the generation of free radicals is higher than the capacity of antioxidant systems to neutralize them, the body enters a state of oxidative stress. Oxidative stress can disrupt redox signaling and control and can cause disruptions in the normal mechanisms of cellular signaling (86).

Inflammation is also often proposed as a contributor to ageing and leukocyte telomere shortening (92, 93). A causal role of inflammation has been suggested in multiple age-associated
diseases including cancer (94). Shorter telomere length is associated with inflammatory markers such as interleukin-6, C-reactive protein, and tumor necrosis factor α (93, 95). Inflammation can accelerate telomere attrition by promoting cell turnover and replicative senescence (96) and by promoting oxidative stress (97). Shorter telomeres are observed in individuals with inflammatory diseases such as ulcerative colitis, liver cirrhosis, chronic kidney disease and chronic obstructive pulmonary disease (93, 98, 99). Just as with oxidative stress various physiological, environmental and lifestyle factors affect inflammation.

2.10 Alcohol Metabolism, Cigarette Smoking, Physical Activity and Oxidative Stress

2.10.1 Alcohol Metabolism

Alcohol induced oxidative stress is linked to the metabolism of ethanol. The effect of ethanol on various tissues is dependent on its concentration in the blood. Blood alcohol concentration is determined by how quickly alcohol is absorbed, distributed, metabolized and excreted by the body (100). Age, sex, diet, smoking, the frequency and the rate of alcohol consumption, presence of food in the stomach and variation in the genes responsible for alcohol metabolism influence blood alcohol concentration (100). Alcohol is metabolized to acetaldehyde by alcohol dehydrogenase (ADH), CYP2E1 from the microsomal ethanol-oxidizing system (MEOS) and to a lesser extent by catalase (100). Each produces free radicals, which are often highly unstable and highly reactive entities that can react with and damage biomolecules such as fats, lipids, and DNA (100). Both acute and chronic alcohol consumption promote the generation of ROS and can interfere with antioxidant defense by decreasing the level or activity of antioxidants (100-102). Additionally, alcohol metabolism in the liver leads to the formation of environments favourable to oxidative stress, including hypoxia, endotoxemia, and the release of
inflammatory molecules that signal inflammation (103, 104). Alcohol can also alter the levels of certain metals in the body, facilitating ROS production (100, 102).

The mitochondrial respiratory chain is the major source of ROS production in the cell (172). Metabolism of alcohol by ADH is coupled to electron and hydrogen atom transfer down the mitochondrial respiratory chain during which a small but significant percentage of oxygen is converted to ROS (105, 106). At the microsomal level, there is an increased generation of oxygen and alcohol derived free radicals, particularly through the action of cytochrome P450 2E1 (CYP2E1). MEOS can also promote oxidative stress indirectly by impairing the defense system against oxidative stress (107). The superoxide radical anion, hydrogen peroxide and, in the presence of iron, the hydroxyl radical are generated by CYP2E1 during alcohol metabolism (107, 108).

2.10.2 Cytochrome P450 2E1 (CYP2E1)

Alcohol metabolism by CYP2E1 results in a significant generation of ROS and contributes considerably to the generation of oxidative stress (171). The induction of CYP2E1 by chronic alcohol consumption has been demonstrated in experimental animals and in humans (109, 110). The regulation of CYP2E1 expression is complex and involves transcriptional, translational and post-translational events (111). Certain genetic differences in CYP2E1 are associated with differences in enzyme activity and enzyme sensitivity to induction. Genetic polymorphisms in the 5’ flanking promoter region of the human CYP2E1 gene have been shown to alter transcriptional regulation (111). The most studied of these is the *5B polymorphism, also known as PstI/RsaI or as the C2 variant depending on the naming convention used. The CYP2E1*5B polymorphism, located in the 5’ flanking region of the CYP2E1 gene, has a minor allele frequency (*5B) of one to five percent in Caucasian populations (112). The *5B variant is associated with higher
transcriptional activity, higher protein levels, and higher enzyme activity than the normal, *1A, form of the gene (113, 114). These functional changes may lead to a greater metabolism of alcohol by CYP2E1 under inducing conditions such as during periods of high alcohol intake and a greater generation of ROS than in individuals with the normal “wild-type” form of the gene. Alcohol has a high affinity for CYP2E1 but the blood alcohol concentration needs to be much higher for alcohol to be preferentially metabolized by CYP2E1 than other alcohol metabolism pathways. However, the alcohol levels required for CYP2E1-mediated metabolism are within the range of concentrations observed in social drinking. The relevance of CYP2E1 in alcohol metabolism increases as blood concentration of alcohol increases. Variation in CYP2E1 genotype may affect the rate of alcohol metabolism and the subsequent generation of oxidative stress, potentially modifying the relationship between alcohol consumption and leukocyte telomere length. Many studies have considered the CYP2E1 genotype-associated risk for cancer and a few have considered an interaction with alcohol on cancer risk. CYP2E1*5B genotype has been linked to increased cancer risk and in some studies has modified the effect of alcohol use on cancer (170).

2.10.3 Cigarette Smoking

Cigarette smoke contains many oxidant molecules (115) and the majority of evidence indicates that cigarette smoking stimulates increased levels of oxidative modifications in DNA (116). Further, smoking appears to have adverse effects on antioxidants. Marangon et al. (117) reported altered plasma antioxidant levels in smokers independent of dietary intake. The presence of a large number of free radicals in cigarette smoke in combination with a greater turnover rate of antioxidant vitamins in smokers than in nonsmokers contributes to a greater generation of oxidative stress in smokers (117). Cigarette smoke can directly damage DNA (120-122) by
interacting with and modifying bases and a greater number of DNA modifications have been observed in lymphocytes of smokers compared to non-smokers (116). Studies have also demonstrated a reduction in the number of DNA modifications with smoking cessation (118, 119). Smokers have on average greater levels of urinary DNA oxidation markers than non-smokers (118, 123, 124, 125) and randomization to smoking cessation results in decreased levels of these markers (118; 126). Numerous studies have confirmed that long-term exposure to cigarette smoke may lead to a systemic oxidant-antioxidant imbalance and chronic low-grade systemic inflammation, indicated by elevated C-reactive protein, fibrinogen, interleukin-6 and increased white blood cell counts (summarized in 123). Additionally, while many smoking-induced changes are reversed after quitting, some inflammatory markers, such as C-reactive protein, may persist 10 to 20 years in former smokers (123). Younger current smokers have lower levels of oxidative stress, as reflected by the ratio of oxidized to total glutathione, than older current smokers indicating an accumulation of oxidative damage with continuing smoking (127).

2.10.4 Physical Activity

An increase in the production of reactive oxygen species is observed in acute exercise; however, this increase acts as a stimulus that up regulates antioxidant enzymes in the longer term. Regular exercise has recognized health benefits and lowers the levels of protein and DNA oxidation products at rest (128). Studies in untrained animals demonstrate increased oxidant levels in response to acute exercise (129, 130) but this effect was counteracted by adaptation to long-term exercise, which induced elevated levels of antioxidant enzymes and reduced free radical production (131-136). These observations from animal models are corroborated by human studies. A cross-sectional study in healthy females demonstrated an increase in antioxidant enzyme activity associated with leisure time physical activity for both low intensity and high
intensity physical activity (137) and several clinical studies have shown that higher strength and physical performance is associated with higher serum levels of antioxidants (138-141). Two intervention studies in sedentary healthy subjects reported increased antioxidant enzymes and resistance to oxidation and decreased inflammation in participants after a 16-week aerobic physical activity (142) and an eight-week standardized aerobic high-amount-high-intensity training program (143) respectively. Elosua et al. (142) reported an increase in antioxidant enzyme activity immediately following acute exercise that lasted at least as long as 24hrs post exercise. Exercise also improves subjective well-being and perceived health status improving overall quality of life (144), which may have indirect effects on a reduction in oxidative stress as a result of reduced psychological stress. Hence, it is possible that the beneficial effects of physical activity on health are due to an adaptive response that increases antioxidant/damage repair enzyme activity and increases resistance to oxidative stress (173).

2.11 The Epidemiology of Alcohol Consumption, Cigarette Smoking, Physical Activity and Telomere Length

2.11.1 Alcohol Consumption

Very few studies have considered the relationship between alcohol consumption and telomere length. Four studies have included alcohol intake as a covariate (147-150) and found no (147, 148, 150) or marginal associations (149) with telomere length. Only two have treated alcohol as the exposure of interest, a case-control study assessing the relationship in a group of men with alcohol abuse (145) and a cohort study in a population of older males, with a mean age of 75 years at the time of telomere length measurement (146). Pavanello et al. (145) reported nearly halved telomeres in men with alcohol abuse relative to those who drank socially. Strandberg et al. (146) reported shorter telomere length in old age with even minor alcohol
consumption in midlife, representing as much as a ten year difference in biological age between zero and highest consumption. However, both studies recruited solely male participants limiting generalizability. Pavanello et al. (145) found no association between number of drink units and telomere length in social drinkers and concluded that it is the condition of alcohol abuse rather than amount of drinking that is associated with shorter telomere length. Strandberg et al. (146) measured telomere length in an older population with a mean age of 76. Telomere length exhibits a higher degree of instability in older populations (12, 55, 151). No studies to date have evaluated the relationship in a mixed gender, young to middle age adult population.

2.11.2 Cigarette Smoking

Six studies were identified with a primary focus on the relationship between age-adjusted telomere length and cigarette smoking and more have included smoking as a covariate. Results from cross-sectional analyses indicate an association with shorter telomere length and a decreasing dose-response effect between pack-years smoking and relative telomere length (76, 149, 152, 153, 154). In longitudinal analysis of the relationship, cigarette smoking in midlife was associated with shorter telomere length and a greater proportion of short telomeres in old age (155); active smoking significantly accelerated the telomere attrition rate in smokers compared to non-smokers (154). Some studies that included smoking as a covariate found a significant inverse association between smoking and telomere length (54, 81, 156-158) but several did not (150, 153, 159, 160, 161, 162, 163). Several of these studies had small sample sizes and may have been underpowered to detect an association.

2.11.3 Physical Activity

Ten epidemiological studies, mostly of cross-sectional design, have been published on the relationship between physical activity and telomere length. Three studies (164-166)
investigated the relationship in long-term endurance-trained athletes and reported an association with longer telomere length compared to moderately active non-athletes (164), sedentary peers (165), and healthy age-matched controls (166). One (165) also observed that the telomere length of older endurance-trained athletes was statistically similar to young endurance-trained athletes, while being longer than that of sedentary peers. These studies suggest an association with telomere length at the extreme end of the physical activity exposure spectrum.

Six studies have been published in non-athlete populations and have produced conflicting results. Two studies, (149,171) reported a lack of association between telomere length and physical activity. One was in an older Chinese population (149), limiting generalizability, and the other (171) in a sample that combined controls with cancer cases. Four studies reported a positive association with telomere length (79, 156, 167, 168) but made different conclusions regarding the intensity of physical activity that confers optimal benefit. Two (156, 167) reported an increasing linear trend in telomere length with greater physical activity and two (79, 168) observed a U-shaped relationship, reporting longer telomere length in moderately active than both inactive and highly active individuals. One additional study reported a moderating effect of exercise on the relationship between stress and telomere length (169). Overall, population studies of physical activity and telomere length have been limited and inconsistent and most studies have considered total activity without considering physical activity by intensity, duration or frequency.

2.12 Minimum Detectable Effects

Power calculations were performed a priori to determine minimum detectable effects for the main variables considered. For the purposes of these calculations telomere length was standardized to a standard normal distribution and had a mean of 0 and standard deviation of 1. The calculations were based on a sample size of 600. Further details of the calculations are
presented in Appendix E. This study had 80% power to detect a difference of 0.40 of a standard deviation in a comparison of ‘high’ alcohol consumption to ‘low’ alcohol consumption and current smoking to never smoking. To improve power alcohol consumption was dichotomized at 7 drinks per week for the purposes of the gene-environment interaction. The detectable difference for effect modification by CYP2E1 genotype was 0.29 of a standard deviation and 0.89 of a standard deviation within wild type and variant strata respectively. The detectable difference estimation for cigarette smoking considered the contrast between current and never smokers. The study had 80% power to detect a difference of 0.34 of a standard deviation. For physical activity the contrast between high and low physical activity levels was considered. The study had 80% power to detect a difference of 0.42 of a standard deviation.

### 2.13 Ethical Consideration

All study participants volunteered to participate in the larger study and provided informed consent. The objectives of this sub-study were consistent with the consent given by participants in the larger cross-sectional study funded by CIHR. Personal identifying information was destroyed and all biological specimens were solely differentiable by laboratory identification numbers. Strict confidentiality procedures were followed regarding all computerized data. All genetic factors in this study do not on their own, pose substantial risk such that identification would warrant further medical inquiry. This project was approved by the Queen’s University Health Science Research Ethics Board (EPID-397-12; see Appendix A).
2.14 Study Objectives and Hypothesis

The objectives of this thesis were:

1. To determine the relationship between level of alcohol consumption and leukocyte telomere length.

   It is hypothesized that higher alcohol intake will be associated with shorter leukocyte telomere length.

2. To determine the interaction between alcohol consumption and CYP2E1 variant CYP2E1*5B.

   The relationship between high alcohol intake and short leukocyte telomere length is hypothesized to be stronger in individuals with the variant *1A/*5B genotype.

3. To determine the relationship between smoking and leukocyte telomere length.

   It is hypothesized that current smokers will have shorter relative leukocyte telomere length than non-smokers. Leukocyte telomere length is expected to decrease with increasing pack-years smoking.

4. To determine the relationship between physical activity and leukocyte telomere length.

   It is hypothesized that physical activity will be associated with longer leukocyte telomere length.
Telomerase in active germ-line cells (eggs and sperm)

Telomerase is repressed in normal somatic cells

Telomerase is reactivated in most cancer cells

**Figure 2-1. Mechanism of telomere shortening due to the end-replication problem.**

Telomerase is active in most germ-line cells and maintains telomere length. Telomerase is repressed in normal somatic cells and telomeres shorten with each cell division. In cancer, telomerase is reactivated and maintains telomeres allowing continued cell division and immortalization.
Figure 2-2. A conceptual diagram of the proposed mechanisms relating selected lifestyle and genetic factors to telomere length and cancer. Stop arrows (red) indicate inhibitory effects while forward arrows (green) indicate promoting events. Alcohol metabolism and cigarette smoke produce reactive oxygen species (ROS), while physical activity promotes increased antioxidative defense. Genetic variation in cytochrome P450 2E1 (CYP2E1), one of the major alcohol metabolizers, has been reported to alter enzyme activity and may affect the extent of ROS generated during alcohol metabolism. CYP2E1 is particularly important at higher alcohol intakes, thus the single nucleotide polymorphism CYP2E1*5B (rs3813867) may modify the relationship with leukocyte telomere length. An imbalance between ROS production and antioxidative defense results in the generation of oxidative stress. Oxidative stress accelerates telomere shortening and promotes inflammation, which may also indirectly accelerate telomere shortening via increased cell replication. Critically short telomeres ordinarily trigger programmed cell death or senescence, but sporadically cells can escape death and senescence and accumulate additional mutations and genetic aberrations resulting in malignant transformation. An even smaller number of these abnormal cells eventually acquire the ability to reactive telomerase initiating cell immortalization and along with other genetic and epigenetic changes cancer progression (174).
References


cessation on oxidative DNA modification estimated by 8-oxo-7,8-dihydro-2'-

127. Nagai K, Betsuyaku T, Kondo T, Nasuhara Y, Nishimura M. Long term smoking with
age builds up excessive oxidative stress in bronchoalveolar lavage fluid. Thorax. 2006;
61:498-503.

128. Powers SK, Ji LL, Leeuwenburgh C. Exercise training-induced alterations in skeletal

129. Leeuwenburgh C, Hansen PA, Holloszy JO, Heinecke JW. Hydroxyl radical generation
during exercise increases mitochondrial protein oxidation and levels of urinary

130. Bejma J, Ji LL. Aging and acute exercise enhance free radical generation in rat skeletal

131. Leeuwenburgh C, Fiebig R, Chandwaney R, Ji LL. Aging and exercise training in

 glutathione antioxidant system to endurance training are tissue and muscle fiber specific.

increases superoxide dismutase activity in ventricular myocardium. Am J Physiol. 1993;
265: 2094.


170. International Agency for Research on Cancer. IARC Monographs Personal Habits and Indoor Combustions Volume 100E: A Review of Human Carcinogens. Consumption of Alcoholic Beverages. Supplementary Web Tables, Section 2, Cancer in Humans. Table 2.77 Studies of CYP2E1 genotype-associated risk for cancer. [Internet]. Lyon France;
2012 [cited 2012 Oct 15]. Available from:

http://monographs.iarc.fr/ENG/Monographs/vol100E/100E-06-Table2.77.pdf


Chapter 3

Measurement of Telomere Length and the Assessment of the Reproducibility of Telomere Length Measures

3.1 Introduction

This study was nested within a larger CIHR funded study that examined the relationship between environmental and lifestyle factors and biomarkers of methionine-homocysteine biosynthesis. The larger study collected a fasting blood sample and questionnaire data from male and female volunteers aged 20 to 50. Stored DNA samples, extracted from whole blood and banked by the larger study, were used for the measurement of telomere length. This chapter describes in detail the optimization of the multiplex qPCR assay for measuring telomere length and the reproducibility of leukocyte telomere length measures beyond that presented in the manuscript in chapter four. Methods specific to the main analysis are presented in manuscript format in chapter four.

A qPCR assay detects and amplifies a targeted region of DNA. It is possible to target a specific region of DNA by designing primer pairs that can interact with and bind to the target area. A DNA polymerase enzyme synthesizes and extends the DNA between the primer pairs. The reaction then repeats this cycle to exponentially amplify the desired region of DNA. A quantification cycle (Cq) value is assigned to the sample by the qPCR reaction machine when the quantity of DNA in the reaction reaches a threshold level. Cq or ‘quantification cycle’ is the metric used for analyzing qPCR results. The Cq represents the qPCR cycle value at which the amplification increases above the threshold level set by the instrument-specific algorithm (1). The threshold is set so that it captures data during the exponential phase of the amplification curve and
is the same for all samples analyzed on a plate. Telomere length was measured using the monochrome, multiplex quantitative polymerase chain reaction (MMQPCR) developed by Cawthon (2).

The MMQPCR assay relies on fluorescence detection by SYBR Green I to monitor the reaction as it progresses. The amount of fluorescence signal generated is proportional to the amount of DNA synthesized during the reaction, which is also dependent on the amount of starting DNA in the reaction (3). The longer the telomere or in other words the more telomere repeats there are the earlier the amplification will be detected. Thus longer telomere lengths have lower Cq values whereas shorter telomeres have higher Cq values. Relative telomere length (rLTL) is estimated as the ratio of telomere repeat copy number to single copy gene copy number (T/S ratio) in each sample relative to the ratio of telomere repeats to single copy gene in a reference DNA sample. This ratio is proportional to the average telomere length of the sample cell population (2).

Differences in the starting concentration of DNA in the reaction can arise during reaction setup; standardization to the single copy gene controls for differences in the amount of starting DNA. Reaction conditions and efficiencies can vary between experimental plates. The T/S ratio is also standardized to an arbitrarily chosen reference DNA sample to account for these differences, allowing for comparability of rLTL between reactions. The telomere length measure obtained by this method is a relative ratio measure of telomere length that should correspond to the relative average telomere length of an individual (2).

The telomere and single copy gene targets are amplified by a different set of primers. To facilitate multiplexing, Cawthon (2) introduced a modification to the single copy gene primers that shifts the melting temperature of the resulting product. The difference in melting temperature
between the telomere product and the single copy gene product ensures that single copy gene signal is measured at a higher temperature after telomere product has melted to baseline and is undetectable. Multiplexing qPCR reactions requires greater optimization; however, multiplexing also increases throughput, decreases sample handling, reduces variation due to experimental setup and minimizes the amount of sample DNA required.

This chapter describes the methods for blood sample acquisition, DNA isolation, and the multiplex qPCR assay as well as the assessment of the reproducibility of telomere length measures. A discussion of method optimization and reproducibility of results is also included.

3.2 Methods

3.2.1 Blood Sample Acquisition

Blood collection devices can interact with blood to alter blood composition, serum, or plasma fractions and affect laboratory tests and other downstream applications (4). A phlebotomist collected blood into serum-separating (SST) and EDTA vacutainer tubes for the larger cross-sectional study. SSTs contain an inert catalyst to facilitate clotting and a gel separator to facilitate serum separation. SSTs were allowed to sit at room temperature for 20-30 minutes, to ensure complete clot formation, prior to serum separation. EDTA tubes contain an anticoagulant that prevents clot formation and were used for whole blood collection. Blood was mixed with anticoagulant by inverting the EDTA tube 8-10 times. Both tubes were labeled with participants’ ID and chilled promptly to preserve integrity. Tubes were centrifuged, within one hour of venipuncture, for 15 minutes in a fixed angle centrifuge at a speed of 3300 rpm for 10 minutes according to manufacturer recommendations. Serum and plasma were separated into test tubes, capped tightly and stored at -80°C until transportation to Ottawa for biochemical analysis. A 250 µL aliquot of whole blood was designated for DNA isolation, transferred to two microfuge tubes,
and stored at -80°C until shipment to the Department of Pathology at Queen’s University. Specimens were shipped according to mandated regulations on the shipment of diagnostic specimens (Transportation of Dangerous Goods Regulations), on dry ice or with frozen ice packs, in sealed Styrofoam boxes, priority, via overnight courier.

3.2.2 DNA Isolation

Extraction serves to purify DNA from inhibitory substances, to stabilize DNA, and to bring DNA to the required concentration for the assay. All of which are required to achieve acceptable sensitivity levels for the qPCR assay. DNA extraction methods typically used to prepare samples for qPCR assays strive to achieve high purities and thus emphasize reducing the presence of inhibitory substances in the sample over maximizing DNA yield. Genomic DNA was purified from leukocytes using the 5-Prime ArchivePure DNA Isolation Kit (Inter Medico, Canada) or the QIAmp DNA Blood Mini Kit (Qiagen, Canada) according to manufactures instructions. Extracted DNA was eluted in Tris-EDTA buffer to protect from degradation, quantified using NanoDrop 2000 UV-Vis spectrophotometer (ThermoScientific, USA) to determine the concentration and purity of the sample, and stored at -20°C to protect from degradation and to maintain stability. The ratio of absorbance at 260 nm to 280 nm on the NanoDrop 2000 UV-Vis spectrophotometer was used to determine the purity of DNA. Generally absorbance ratios of 1.8 - 2.0 indicate pure DNA (5). The range of absorbance ratios for the samples used in this analysis was 1.70-1.97.

3.2.3 Multiplex qPCR Assay for Measuring Telomere Length

The reference DNA sample chosen for this study was a sample of pooled DNA from a 29 year-old female and a 49 year-old male and was strategically selected so that it might represent the middle of the range of values obtained in the study population. The T/S ratio is equal to 1
when the sample DNA is identical to the reference DNA in the ratio of telomere repeat copy number to single copy gene copy number (2). The age range of the study population was 20 to 50 years. Based on the selection of the reference sample it was expected that the mean rLTL in this study would be close to 1. Human betaglobin was chosen as the single-copy gene for this assay as this primer set was modified for multiplexing, tested and previously published by Cawthon (2).

Reaction components and final concentrations are presented in Appendix B. Samples were run on 96-well plates using the BioRad CFX96 PCR Detection System set to the thermal cycling profile described in Appendix B. A 3-fold serial dilution of a pooled DNA sample was used to prepare a 5-point standard curve ranging from 150 ng to 1.85 ng of DNA. Reaction efficiencies are calculated from the standard curve by the BioRad software using the linear regression slope of the dilution series and the equation $E=10(-1/slope)$. Reaction efficiencies are used to calculate relative telomere length based on an adaptation of the Pfaffl method (6) according to the equation presented below (7):

$$ T \over S \text{ ratio} = \frac{1 + E_{Tel}C_{Tel}}{(1 + E_{Scg})C_{Scg}} $$

Relative $ T \over S \text{ ratio} = \frac{\text{Mean } T \over S \text{ ratio}_{sample}}{\text{Mean } T \over S \text{ ratio}_{reference}}$

where $E = \text{Efficiency of the qPCR reaction calculated using the linear regression slope of the standard curve dilution series} = 10(-1/slope)$; $Tel = \text{Telomere reaction}$; $Scg = \text{Single copy gene (human $betaglobin$) reaction}$; $Cq = \text{Quantification cycle, the fractional cycle number where fluorescence increases above the threshold}$. 
3.2.4 Rejection Criteria for qPCR Samples

The ideal slope for a standard curve dilution series is -3.32, which correlates with a reaction efficiency of 100%. An efficiency of 100% indicates perfect doubling in the amount of PCR product during exponential amplification and is the best indicator of a robust, reproducible assay (8). Slopes in the range of -3.60 to -3.10, which correlate to efficiencies between 90 and 110%, are generally considered acceptable for qPCR (8). Higher efficiencies indicate the presence of inhibitors in the reaction that may delay the Cq in samples with the highest concentrations. Efficiencies below 90% indicate poor reaction conditions or errors during experimental set-up. The R² of the slope of the standard curve indicates how well the experimental data fit the line. A significant difference in Cq values between replicates will lower the R² value. An acceptable R² for the standard curve linear regression is > 0.98. Plates with efficiencies outside the range of 90-110% and/or an R² value below 0.98 were rejected and all samples on the plate were redone. Samples with a qPCR signal that amplified outside the range of the standard curve and samples with a CV between triplicates that was greater than 10% were considered unacceptable and were redone. Samples that were redone three times and failed to meet the acceptability criteria described above were excluded from analysis.

3.2.5 Analysis Strategy for Assessing Reproducibility of Telomere Length Laboratory Measures

Each sample was measured in triplicate and the repeated-measure values were used to assess reproducibility of telomere length measures. The coefficient of variation (CV) and the intraclass correlation coefficient (ICC) were calculated to estimate the percentage of the total variation that was attributable to measurement error. The CV measures the extent of the variability in relation to the mean and is defined as the percent ratio of the within-person standard
deviation to the within-person mean. Both the inter-assay and the intra-assay CVs were estimated from a random-effects one-way analysis of variance (ANOVA) model. ANOVA tests the null hypothesis that the between-person variation is equal to zero against the alternative hypothesis that the between-person variation is greater than zero. Under the null hypothesis all of the variation observed between study participants is attributable to within-person variation or ‘noise’ while under the alternative hypothesis there is a true underlying difference between means for individuals.

The CV was estimated from ANOVA by dividing the root mean square error by the grand mean of telomere length. Generally, CVs under 20% are considered desirable, whereas CVs greater than 30% indicate unstable results (9).

\[
CV = 100\% \times \frac{\text{within-person standard deviation}}{\text{within-person mean}}
\]

The ICC also measures the reproducibility of replicate measures from the same subject and is defined as the between-person variance divided by the sum of the between-person and the within-person variance (14). An ICC less than 0.4 indicates poor reproducibility, an ICC between 0.4 and 0.75 indicates fair to good reproducibility, and an ICC greater than 0.75 indicates excellent reproducibility (10). The ICC was estimated by one-way ANOVA.

\[
\begin{align*}
 \eta_i &= \frac{mSS_B - SS_T}{(m-1)SS_T} \quad \text{with a two - sided } 100\% \times (1 - \alpha) \ CI : \\
 c_1 &= \max \left\{ \frac{F}{F_{k-1,N-k,1-\alpha/2}^{-1}} : \frac{F}{n_0 + \frac{1}{n_0+k-k_1-1}} \right\} \\
 c_2 &= \max \left\{ \frac{F}{F_{k-1,N-k,1-\alpha/2}^{-1}} : \frac{F}{n_0 + \frac{1}{n_0+k-k_1-1}} \right\} \\
\end{align*}
\]

where \( \eta_i \) = number of observations per subject; \( SS_B \) = sum of squares between subjects; \( SS_T \) = corrected total sum of squares.
3.3 Results

According to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (5), a melt curve analysis or a gel are both acceptable methods for demonstrating primer specificity. Figure 3-1 shows the melt curve for the multiplex reaction, with two peaks, the first for the telomere and the second for the single copy gene product, which melt at different temperatures. The melt curve also demonstrates that the single copy gene signal is acquired when the telomere product has completely melted. A gel was also run to confirm that a single telomere and single single-copy gene product were generated at the expected product size of 79 bp and 106 bp respectively (Figure 3-2).

The standard curves in Figure 3-3 serve as an example of acceptable standard curves for the beta-globin and telomere reactions. The log plots of DNA concentration versus cycle threshold show that the curves are linear over the 81-fold DNA concentration range. The reaction efficiencies and $R^2$ are printed below the curve. The reaction efficiencies for both reactions are within the 90-110% range and the $R^2$ also meet the $>0.98$ cutoff. The same pooled-DNA sample was used to generate standard curves for every plate. Standard curve efficiencies were used to calculate relative T/S ratios.

Table 3-1 summarizes the frequency and percent of samples for which a measure of relative telomere length could not be obtained stratified by age category. A reliable measure of telomere length could not be obtained for approximately 5% of the total sample size, or 25 out of 529 samples.

The statistical reproducibility of telomere length measurement was determined using the coefficient of variation and the intra-class correlation coefficient calculated from one-way analysis of variance. Two aspects of the reproducibility were assessed: the intra-assay
reproducibility of three replicates for each sample on a plate and the inter-assay reproducibility of mean telomere length for ten samples included on different plates (Table 3-2). The intra-assay CV of 8.98% was calculated on 511 samples with available DNA and an obtained measure of rLTL. The inter-assay CV of 6.39% was calculated on ten samples, chosen to represent the age and sex distribution of the study population, that were included on three different plates over three different days, two on the same day and one on a later day. The ICC was 0.99 with a 95% confidence interval of 0.989 to 0.991.

3.4 Discussion

Quantitative qPCR has become a common tool because of its speed and ease of scalability for high throughput analysis. However, the quality of data generated by qPCR is highly dependent on careful optimization of the assay. The tested aspects of the assay included primer specificity, good reaction efficiency, and an adequate quantitative range.

Primer specificity was demonstrated by melt curve analysis and gel. The telomere and betaglobin primers generated products that melted at the expected melting temperatures and were of the expected size. From the gel the no template control (NTC) shows a small amount of primer dimer formation. However, this is an inherent part of the MMQPCR assay design. A telomere primer dimer product always appears in the NTC reaction wells at high cycle numbers. This does not matter as long as the NTC signals are still at baseline over the cycle range of the Cq values for the standard curve and all of the sample Cq values are collected over a cycle range where again the NTC signal is still at baseline (personal communication, Richard Cawthon, 15).

Standard curves were used to determine reaction efficiencies, R², and to confirm that samples amplified within the linear dynamic range of the standard curve. Efficiencies for all reactions were in the range of 90-110% and the R² for all plates was greater than 0.98. All plates
with values outside this range were redone. The reaction efficiency, $R^2$, and shape of the amplification curve are indicators of accurate sample and reagent pipetting as well as a reaction that is well optimized and does not contain inhibitory compounds (1). Relative telomere length can only be accurately estimated from the standard curves if the sample amplifies within the linear dynamic range of the curve. To avoid errors in quantification all samples that amplified outside the dynamic range of the standard curve were not included in the final analysis. Reliable estimates of telomere length were obtained for 95% of samples with available DNA.

Results from Table 3-1 suggest that the MMQPCR assay did not systematically fail for the oldest or the youngest study participants, those with potentially the shortest or the longest telomeres respectively. These samples may have had carry-over contamination and reaction inhibitors that impeded accurate quantification (1).

A coefficient of variation and the interclass correlation coefficient were calculated to assess the reproducibility of telomere length measurement. The coefficient of variation was also calculated to estimate within-plate and between-plate reproducibility and both were well below the desirable limit of 20% (9, 13). The intra-assay CV calculated by Lan et al. (11), on telomere length measurement performed in the laboratory of Cawthon (2), was 11%. In his original publication Cawthon (2) reported an intra-assay CV of 5.22%, others have reported inter-assay CVs in the range of 2.27 to 28% (12). The ICC of 0.99 indicates a high degree of reproducibility between triplicate measures of the same sample.

The relatively low CVs and high ICC for this study suggest that the error introduced by the measurement method for relative telomere length was small enough that meaningful differences could be detected between groups and that the measurement error would not
substantially reduce the statistical power for subsequent analysis. Telomere length values were therefore deemed to be suitable for use in the hypothesis tests of the present study.
3.5 Figures and Table

Figure 3-1. Multiplex qPCR assay specificity verification by melt curve analysis showing telomere and betaglobin products. The telomere (79°C) and betaglobin product (91°C) melt at different temperatures with an approximately 12°C difference between the melting temperatures of the two products.
Figure 3-2. Gel showing the single copy gene (hbg) and telomere (tel) products. The single copy gene and telomere products were the expected size of 106 and 79 bp respectively. A 50 bp DNA ladder was included as a reference. A 6% TBE agarose gel was prepared to generate this image. Abbreviations: NTC: No template control; S1: sample 1; S2: sample 2; S3: sample 3; Hbg: human betaglobin; Tel: telomere; bp: base pair
A. Betaglobin

B. Telomere

Figure 3-3. Example of standard curves for betaglobin and telomere reactions. Reaction efficiencies are presented below the curve and are used in the calculation of relative telomere length. Panel A shows the standard curve for the betaglobin reaction. Panel B shows the standard curve for the telomere reaction. Standard curves were prepared using the same pooled-DNA sample for every plate in the study.
Table 3-1. Distribution of samples with no amplification or weak qPCR signal by age category.

<table>
<thead>
<tr>
<th>Age category</th>
<th>Frequency</th>
<th>No amplification /weak qPCR signal</th>
<th>Percent of age category</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 - 25</td>
<td>136</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>26 - 30</td>
<td>82</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>31 - 35</td>
<td>76</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>36 - 40</td>
<td>77</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>41 - 45</td>
<td>89</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>46 - 50</td>
<td>69</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>529</strong></td>
<td><strong>25</strong></td>
<td><strong>5</strong></td>
</tr>
</tbody>
</table>
Table 3-2. Reproducibility of telomere length measurement calculated using one-way analysis of variance.

<table>
<thead>
<tr>
<th>Test</th>
<th>Samples</th>
<th>Root MSE</th>
<th>Grand Mean</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-assay CV&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10</td>
<td>0.044</td>
<td>0.69</td>
<td>6.39 %</td>
</tr>
<tr>
<td>Intra-assay CV&lt;sup&gt;b&lt;/sup&gt;</td>
<td>511</td>
<td>7.42</td>
<td>82.71</td>
<td>8.98 %</td>
</tr>
<tr>
<td>ICC&lt;sup&gt;b, c&lt;/sup&gt;</td>
<td>511</td>
<td>-</td>
<td>-</td>
<td>0.990 (0.989, 0.991)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Abbreviations: MSE=mean square error; CV=coefficient of variation; ICC=intraclass correlation coefficient

<sup>a</sup> Mean rLTL of 3 replicates for each of 10 samples was compared between 3 plates

<sup>b</sup> For each sample 3 replicates were included on each plate

<sup>c</sup> ICC calculated using sum of squares (model) = 9681206.14, sum of squares (corrected total) = 9735486.01, and m=3

<sup>d</sup> ICC (95% confidence interval)
References

1. Basic principles of qPCR. [Internet]. USA: Thermo Scientific; [cited 29 May 2014]


Chapter 4

Alcohol Consumption, Cigarette Smoking, and Physical Activity in Relation to Leukocyte Telomere Length

4.1 Abstract

4.1.1 Background

Telomeres protect chromosome ends from degradation during DNA replication and are important in maintaining chromosomal integrity and stability. Epidemiologic studies support the relationship between shorter telomeres and risk of cancer at several sites. However, there is limited knowledge about the lifestyle determinants of telomere length. The objective of this study was to determine the effect of three lifestyle factors known to be important in cancer etiology, on relative leukocyte telomere length (rLTL): alcohol consumption, smoking, and physical activity.

4.1.2 Methods

This cross-sectional study included 477 healthy male and female volunteers aged 20 to 50 years. Subjects completed a questionnaire and provided a fasting blood sample for biomarker and biochemical analysis. We used multiplex quantitative real-time PCR (qPCR) to measure rLTL. The relationship between the lifestyle exposures of interest and rLTL was assessed using multiple linear regression while controlling for important covariates.

4.1.3 Results

There was no association between alcohol consumption and rLTL. Current smokers and those in the middle and lower tertile of pack-years smoking had on average shorter rLTL than never smokers (p=0.02). A linear trend of longer rLTL was observed with increasing total
physical activity (p=0.06). Additionally, compared to the lowest quartile, being in the highest quartile of vigorous physical activity was associated with on average longer rLTL. A significant linear trend of increasing rLTL with increasing MET-minutes of vigorous physical activity was observed (p=0.02).

4.1.4 Conclusions

This study found that lifestyle factors, specifically cigarette smoking and vigorous physical activity have an impact on telomere length. Smoking was related to shorter telomere length while vigorous physical activity was related to longer telomeres.

4.2 Introduction

Telomeres are repeating “TTAGGG” DNA sequences located at the ends of linear chromosomes. They protect chromosomes from damage and consequently are important for maintaining chromosomal stability (1). As replicative machinery cannot completely copy the ends of linear chromosomes, 50-100 base pairs of telomeric DNA are lost at each cell division; this results in gradual telomere attrition with increasing age (2). In recent years, research has demonstrated a relationship between short telomeres and ageing related diseases, including many forms of cancer. Evidence from two recent meta-analyses supports an inverse association between telomere length and risk of cancer with the strongest evidence existing for bladder, esophageal, gastric, and renal cancers (11,12).

Critically short telomeres ordinarily trigger either cell death or replicative senescence (3,4). However, rarely cells may bypass cell death or senescence and continue to divide acquiring mutations and genetic aberrations that may eventually result in malignant transformation (5, 6). Replicative senescence may indirectly promote cancer initiation by contributing to
immunosenescence (7), an age-dependent decrease in immunological functioning, as well as a lack of immunosurveillance (8), the body’s ability to monitor and recognize aberrant cells, which are associated with the development of aging-related diseases such as cancer (9, 10).

Due to the nature of their molecular structures telomeres are particularly vulnerable to oxidative stress (13). DNA damage that occurs as a result of oxidative stress accelerates telomere shortening. It has been well documented that oxidative stress is associated with shorter telomeres (14-18). In addition, oxidative stress often accompanies inflammation, and can contribute to leukocyte telomere shortening by promoting increased cell turnover and replicative senescence (19-21).

Various physiological, environmental and lifestyle factors can affect oxidative stress and inflammation, and presumably telomere length. Smoking (22, 23) and alcohol consumption (24, 25) lead to an increase in oxidative damage and inflammation. Both chronic and acute alcohol consumption promote increased production of reactive oxygen species (ROS) (28, 29) and enhanced peroxidation of lipids, proteins and DNA (30). Cigarette smoke contains many oxidants and free radicals (31) that can both directly and indirectly cause oxidative damage to DNA (32). Regular physical activity is associated with decreased levels of oxidative stress and inflammation and is known to confer many health benefits (26, 27). Physical activity is associated with increased antioxidant and enzyme activity levels (27, 33) and has been shown to regulate telomere-stabilizing proteins (34).

The objective of this cross-sectional study was to determine the relationship between alcohol consumption, smoking and physical activity and relative leukocyte telomere length.
4.3 Materials and Methods

4.3.1 Design and Source Population

This analysis is nested within a cross-sectional study that recruited 678 healthy male and female volunteers, aged 20 to 50 years. Volunteers were recruited from Ottawa, Ontario, Kingston, Ontario and Halifax, Nova Scotia. Complete study design details and subject eligibility and exclusion criteria have been described elsewhere (35). Briefly, participants were recruited between 2006 and 2008 by informed consent in approximately equal numbers by age and sex. Each volunteer was asked to complete a self-administered research questionnaire and provide a fasting blood sample for biochemical and genetic analysis. Ethics approval for the study was obtained from the Queen’s University Health Science Research Ethics Board. Of the 678 recruited participants 80 were excluded from this study because of an inadequate blood sample, 69 were excluded due to insufficient DNA, a further 25 had qPCRs that produced product amounts outside the dynamic range of the standard curve, and an additional 28 subjects were excluded because of missing covariate information. The final sample size for this analysis consisted of 477 individuals with available DNA and complete questionnaire and biochemical data.

4.3.2 Exposure Measurement

Demographic and lifestyle information, including alcohol consumption, cigarette smoking, physical activity and information on other possible risk factors, was assessed by questionnaire. Alcohol consumption was estimated using validated global quantity-frequency questions adapted from the 1996 Ontario Drug Monitor (ODM). The definition of one standard alcoholic beverage (1 alcoholic beverage = 12oz beer (1 bottle), 5oz wine (1 glass), or 1.5oz liquor) was provided in the questionnaire. Information was collected on the typical number of
drinking occasions over the past month and the typical number of drinks consumed per drinking occasion. Frequency and quantity were combined to estimate weekly alcohol consumption according to recommendations from the ODM Technical Guide (36). Alcohol consumption was categorized into four levels, ‘abstainer’, ‘low’, ‘moderate’ and ‘high’, according to published low-risk drinking guidelines (38). The ‘abstainer’ category included both lifetime abstainers and former drinkers and was defined as 0 drinks/week. ‘Low’ consumption was defined as an intake of more than 0 but fewer than 7 drinks/week while ‘moderate’ and ‘high’ consumption had sex specific category definitions. Females consuming more than 7 but fewer than 10 drinks/week and males consuming more than 7 but fewer than 15 drinks/week were classified as ‘moderate’. ‘High’ consumption was defined as greater than 10 drinks/week for females and greater than 15 drinks/week for males.

Smoking status was categorized as ‘never’, ‘former’, and ‘current’. Participants who reported current, daily smoking of at least one cigarette per day were classified as ‘current’ smokers. Those who were not currently smoking, but reported smoking at least one cigarette a day for at least six months at some point in their lives were classified as ‘former’ smokers. Otherwise, participants were coded as ‘never’ smokers. For the purpose of the analysis, packs per day were calculated assuming a 20-cigarette pack. For current smokers, smoking duration was calculated from the age smoking started to age at the time of questionnaire completion and for former smokers, age started smoking to age stopped smoking. Cumulative smoke exposure was calculated as the product of the number of packs smoked per day and smoking duration. Pack-years were categorized into tertiles. All variables were coded as indicator variables that used never smokers as the reference category.
Physical activity was measured using the short-form International Physical Activity Questionnaire (IPAQ), which was developed and validated for use in adults aged 15 to 69 years (39). The IPAQ short form instrument measures walking, as well as moderate-intensity and vigorous-intensity activities across four domains: leisure-time physical activity, domestic and gardening activities, work-related physical activity, and transport-related physical activity. Total physical activity MET-minute scores were calculated by summing duration, in minutes, and frequency, in days, of walking, moderate-intensity, and vigorous-intensity activities. The IPAQ scoring protocol (40) was used to guide data cleaning and categorization of total physical activity into ‘low’, ‘moderate’ and ‘high’ categories. Moderate activity was defined by the following criteria: at least 20 minutes of vigorous activity per day on three or more days per week; five or more days of moderate-intensity activity and/or walking of at least 30 minutes per day; or five or more days of any combination of walking, moderate-intensity or vigorous-intensity activities achieving a minimum of at least 600 MET-minutes/week. ‘High’ activity was defined as vigorous-intensity activity on at least 3 days and accumulating at least 1500 MET-minutes/week; or seven or more days of any combination of walking, moderate- or vigorous-intensity activities equivalent to at least 3000 MET-minutes/week. ‘Low’ activity was defined as either no activity reported or any amount of activity that did not meet the criteria for moderate or high. Vigorous physical activity was considered as a separate variable. Vigorous physical activity level was calculated according to the IPAQ continuous score protocol (40), by multiplying 8 METs by the number of minutes of vigorous physical activity per day and by the number of vigorous days per week of vigorous physical activity; finally, this variable was categorized into quartiles.
4.3.3 Measurement of Relative Leukocyte Telomere Length

Blood sampling, processing, and aliquoting procedures are described in detail elsewhere (35). A whole blood sample was obtained from each study participants at enrollment and stored at -80°C until DNA extraction. Genomic DNA was isolated from leukocytes, according to manufacturer’s instructions, using the 5Prime ArchivePure DNA Blood Kit or the QIAmp DNA Blood Mini Kit. Extracted DNA was eluted in TE buffer. The concentration and purity of each sample was measured with the NanoDrop 2000 UV-Vis spectrophotometer and stored at -20°C until analysis.

rLTL was measured by monochrome multiplex qPCR described by Cawthon (41). This assay measures the relative ratio of telomere repeat copy number (T) to single gene copy number (S) (human betaglobin) in experimental samples compared to a reference DNA sample, producing a relative T/S ratio measure. Twenty nanograms of template DNA diluted in 10 µl of molecular grade water (VWR) were added to 15 µl of master mix, prepared as reported by Cawthon (41) for a final reaction volume of 25 µl. Primer oligonucleotides were manufactured by Integrated DNA Technologies per sequence descriptions reported by Cawthon (41). Samples from all individuals were assayed in triplicate on 96-well plates using the BioRad CFX96 Real-Time PCR Detection System set to the thermal cycling profile specified by Lan et al. (42). Five concentrations (3 fold serial dilution: 150, 50, 16.7, 5.5 and 1.9 ng) of a pooled DNA sample were included on each plate, in triplicate, providing data for the generation of standard curves. Standard curve efficiencies ranged from 90 to 110% and were used in relative quantification of rLTL by the Pfaffl method (43). rLTL for each sample was standardized to a reference sample prepared from pooled DNA from a 29-year-old female and 49-year-old male. Statistical tests of reliability were calculated from a random-effects one-way analysis of variance model using mean square error as
an estimate of within-person standard deviation. The intraclass correlation coefficient for triplicate measures was 0.990 and the intra-assay coefficient of variation was 8.98%. The inter-assay coefficient of variation of 6.39% was calculated on ten quality control samples run on three different plates. Samples that did not amplify or generated a signal outside the dynamic range of the standard curve were excluded from analysis. A measure of rLTL could not be obtained for 4.7% of the samples with available DNA; however, these samples did not consistently belong to any specific age group suggesting that the method did not systematically fail for the longest or the shortest telomere length.

### 4.3.4 Statistical Analysis

Multivariable linear regression was used to determine regression coefficients for the relation between alcohol consumption, smoking, and physical activity variables and rLTL. Exposure metrics for smoking and physical activity included smoking status, tertiles of cumulative pack-years smoking, total physical activity, and quartiles of vigorous physical activity. Potential covariates included age, sex, ethnicity, BMI, and lipid ratio. A parsimonious model predicting rLTL from among these covariates was selected using backward elimination with a liberal p-value of 0.2 and alcohol, smoking, and physical activity forced in the model. The final parsimonious model included age, sex, and lipid ratio and was used to adjust the analyses of the lifestyle exposures of interest. Participants were recruited at study centres based in three different municipalities; thus a random-effects parameter was included in the model to account for cluster sampling. A p-value for trend across ordered categorical variables was calculated by
assigning integer values to categories and assessing the parameter as a continuous variable (69).

All statistical tests were two-tailed and conducted in SAS 9.3 (SAS Institute, Cary, NC).

4.4 Results

A summary rLTL distribution and the covariates considered are presented in table 4-1. rLTL was approximately normally distributed with a right skew, mean of 0.85 and standard deviation of 0.30. Subjects where aged 20 to 50 with an approximately univariate distribution. Age was inversely associated with rLTL; every 10-year increase in age was associated with an rLTL decrease of 0.060 relative T/S ratio units (p<0.01). The study population was 57% female, predominantly Caucasian, and 44% of the study population were overweight or obese. Age-adjusted effects are presented for the covariates considered; none were predictive of rLTL in this analysis. Age, sex, and lipid ratio met the inclusion criteria for the parsimonious model and the analysis of the lifestyle exposures of interest controls for this set of covariates.

Table 4-2 contains results of multivariable linear regression with age-adjusted and fully adjusted effects presented. This description of results focuses on the fully adjusted results. In multivariable linear regression, alcohol consumption was not related to rLTL (p=0.57). Compared to never smoking, current smoking was associated with shorter rLTL (p=0.02). Pack-years smoking was a significant predictor of rLTL (p<0.01) and those with cumulative pack-year smoking in the lower and middle tertile had shorter rLTL than never smokers. The coefficient estimates for current smoking indicate that, on average, rLTL was 0.096 relative T/S ratio units shorter for current smokers than never smokers. Being in the lower and middle tertile of pack-years smoking was associated with, on average, 0.089 and 0.14 relative T/S ratio unit shorter rLTL compared to never smokers.
While total physical activity was not a significant predictor in the model, there was evidence of a positive dose-response relationship with rLTL across categories of total physical activity ($p_{\text{trend}}=0.06$; Table 4-2). Several studies have reported longer telomeres in endurance-trained athletes; thus, we considered vigorous physical activity separately from total physical activity. Vigorous physical activity in the fourth, relative to the first, quartile was associated with longer telomere length ($p<0.01$). rLTL was 0.12 relative T/S ratio units longer for those in the fourth quartile of vigorous physical activity compared to the first and there was a significant linear trend of increasing rLTL with greater vigorous physical activity ($p=0.01$).

4.5 Discussion

An association of rLTL with alcohol consumption was not observed in this population of healthy 20 to 50 year old men and women. However, rLTL was inversely associated with current smoking and pack-years smoked. A linear trend of longer telomere length was observed with increasing total physical activity, although this trend was not statistically significant. Vigorous physical activity was positively associated with rLTL.

Oxidative stress and inflammation are known to accelerate age-dependent telomere shortening and likely mediate the association between lifestyle exposures and rLTL. Oxidative stress causes direct and indirect damage to DNA and under a state of high oxidative stress accumulation of single-strand breaks is the major cause of telomere shortening (16, 44). Telomeres are particularly sensitive to base oxidation by ROS because of the large number of guanine (G) nucleotides in their repeating TTAGGG sequences (44, 45). Environmental and lifestyle factors can act to either promote or mitigate oxidative stress. ROS are generated as a byproduct of alcohol metabolism (28) and cigarette smoke (31, 32); in addition, both alcohol and cigarette smoke can interfere with antioxidative defense (23, 25). Regular physical activity
augments the physiological antioxidant response (26, 27). Additionally, physical activity
promotes energy expenditure and helps maintain energy balance, thus being physically active
may reduce obesity (46) and the associated levels of oxidative stress and inflammation (47). The
extent of ROS damage to telomeric DNA depends on a balance between factors that promote
oxidative stress and those that augment antioxidative defense.

While moderate alcohol consumption, such as that observed in our study, does not appear
to affect telomere length, evidence from other studies suggests that heavy drinking accelerates
telemore shortening. Our findings differed from the results of two other studies with a primary
focus on the relationship between alcohol and rLTL. Contrary to our observations, both reported
an association between alcohol intake and rLTL; one in a cross-sectional analysis in a group of
males with alcohol abuse (48) and the other in longitudinal analysis in a group of older
businessmen (30). The range of alcohol consumption levels reported by these studies was well
above that seen in our population of healthy volunteers. The reported association of alcohol
consumption with rLTL was observed only with alcohol intakes of above 28 drinks/week.

It is also possible that the time of exposure measurement, at healthy midlife, by
Strandberg et al. (30), represents a more biologically relevant and accurate exposure measure for
the relationship between alcohol consumption and rLTL. Alcohol use at this age is more stable
and is likely more representative of a lifetime cumulative exposure.

Our observation of a relationship of rLTL with current smoking and pack-years smoking
is consistent with the existing literature. Seven other studies have demonstrated an inverse
association between smoking (49-55) and rLTL; one study reported a threefold increased yearly
attrition rate in smokers compared to non-smokers (54).
Our results corroborate reports of a positive linear trend of increasing rLTL with total physical activity and vigorous physical activity (56, 57). However, some studies have reported a U-shaped relationship between physical activity and telomere length and found that moderate physical activity was associated with longer telomere length than both inactivity and vigorous physical activity (58, 59). Others have reported longer telomere length in endurance-trained athletes relative to both moderately active non-athletes (60) and sedentary peers (61). In addition to promoting increased antioxidant activity, greater physical activity may indirectly protect rLTL by moderating the effect of psychological stress (62). Both short and long-term exercise has been shown to elevate telomerase activity (34, 63, 64).

To facilitate meaningful interpretation of these results our findings are presented in the context of several limitations. This was a cross-sectional analysis; therefore the associations observed are not necessarily causal. Our study population was 20 to 50 years old and primarily Caucasian; thus, generalizability to other populations may be limited.

While the monochrome multiplex qPCR method for measuring telomere length is commonly used in epidemiologic studies and is scalable for high throughput analysis (65,66), it measures mean leukocyte telomere length in relative T/S ratio units and does not yield absolute values or percentage of short telomeres. Leukocytes consist of a mixture of cell subsets with different individual telomere lengths; however, there is strong correlation in telomere length between cell subsets within an individual (67). Additionally, the variation in qPCR measurement is greater than that for the Southern blot method of measuring telomeres (68). However, while Southern blot is more precise, it requires large amounts of DNA and cannot be performed on DNA with even modest degradation (68).
Considering that rLTL shortens overtime, it may be important to measure cumulative exposures. Our measures of the lifestyle exposures of interest were self-reported and assessed over a one month recall period that may not represent a cumulative lifetime dose. To assess the effect of these lifestyle exposures directly on telomere attrition a prospective study, that measures telomere length at multiple time points, is required. Additionally, the potential for unmeasured or residual confounding exists as we have insufficient information about certain factors that are associated with healthy ageing and may be related to telomere length.

In conclusion, in this study relative leukocyte telomere length was associated with vigorous physical activity and smoking, but was not associated with alcohol consumption. The results suggest that lifestyle factors may influence telomere dynamics and leukocyte telomere shortening. Longitudinal studies that measure telomere length at multiple time points and adjust for important environmental and genetic covariates are needed to confirm these results.
4.6 Acknowledgements

The authors would like to thank the study participants and the study coordinator Gwyneth Fairfield. The authors would also like to thank Raico Ernesto Laria Lamela for his contributions to the optimization of the telomere length assay and Sean Taylor for his technical support.

4.7 Funding

This work was supported by an operating grant from the Canadian Institutes for Health Research (CIHR). L. Latifovic was supported in part by the Terry Fox Foundation Transdisciplinary Training program in Cancer Research in partnership with the Canadian Institutes for Health Research (CIHR).

4.8 Disclosure of Potential Conflicts of Interest

No potential conflicts of interest to disclose.

4.9 Author’s Contributions

Conception and design: L. Latifovic and W.D. King
Development of methodology: S.D. Peacock, L. Latifovic, W.D. King, T.E. Massey
Acquisition of data: L. Latifovic, S.D. Peacock, W.D. King
Analysis and interpretation of data: L. Latifovic
Drafting of manuscript: L. Latifovic
Review and/or revision of manuscript for important intellectual content: W.D. King, T.E. Massey, S.D. Peacock
Administrative, technical, or material support: S.D. Peacock
Study supervision: W.D. King and T.E. Massey
### 4.10 Tables

**Table 4-1. Characteristics of the study population and age-adjusted rLTL.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N (%)</th>
<th>Coefficient ± SE</th>
<th>p-value a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-29</td>
<td>192 (40)</td>
<td>Ref.</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>30-39</td>
<td>138 (30)</td>
<td>-0.061 ± 0.032</td>
<td></td>
</tr>
<tr>
<td>40-50</td>
<td>147 (31)</td>
<td>-0.13 ± 0.032**</td>
<td></td>
</tr>
<tr>
<td><strong>Age (per 10 years)</strong></td>
<td></td>
<td>-0.060 ± 0.015</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td>Male</td>
<td>205 (43)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>272 (57)</td>
<td>-0.038 ± 0.027</td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td>0.35</td>
</tr>
<tr>
<td>Caucasian</td>
<td>409 (86)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>33 (7)</td>
<td>0.046 ± 0.053</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>35 (7)</td>
<td>-0.054 ± 0.051</td>
<td></td>
</tr>
<tr>
<td><strong>Body mass index (kg/m²)</strong></td>
<td></td>
<td></td>
<td>0.53</td>
</tr>
<tr>
<td>Normal</td>
<td>266 (56)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>144 (30)</td>
<td>-0.022 ± 0.030</td>
<td></td>
</tr>
<tr>
<td>Obese</td>
<td>65 (14)</td>
<td>-0.042 ± 0.040</td>
<td></td>
</tr>
<tr>
<td><strong>p-trend</strong></td>
<td>0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lipid ratio (mmol/L)</strong></td>
<td></td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>Q1</td>
<td>122 (26)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>116 (24)</td>
<td>0.004 ± 0.037</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>120 (25)</td>
<td>-0.062 ± 0.037</td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>119 (25)</td>
<td>0.012 ± 0.039</td>
<td></td>
</tr>
<tr>
<td><strong>p-trend</strong></td>
<td>0.82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** SE=standard error; Ref.=referent

a Calculated by F-statistic from random-effects ANCOVA adjusted for age and including a random-effects parameter representing study center; except for age (per 10 years) which was calculated from simple linear regression.

**p<0.01**
Table 4-2. Results from multivariable linear regression presenting adjusted regression coefficients relating lifestyle exposure variables to rLTL.

<table>
<thead>
<tr>
<th>Lifestyle exposure</th>
<th>Age Adjusted (^a)</th>
<th>Fully Adjusted (^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>Coefficient ± SE</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abstainer</td>
<td>45 (9)</td>
<td>-0.006 ± 0.046</td>
</tr>
<tr>
<td>Low</td>
<td>325 (68)</td>
<td>Ref.</td>
</tr>
<tr>
<td>Moderate</td>
<td>68 (14)</td>
<td>-0.046 ± 0.038</td>
</tr>
<tr>
<td>High</td>
<td>39 (8)</td>
<td>-0.025 ± 0.049</td>
</tr>
<tr>
<td>(p_{trend})</td>
<td></td>
<td>0.39</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>325 (68)</td>
<td>Ref.</td>
</tr>
<tr>
<td>Former</td>
<td>90 (19)</td>
<td>(-0.078 ± 0.035^*)</td>
</tr>
<tr>
<td>Current</td>
<td>62 (13)</td>
<td>(-0.094 ± 0.040^*)</td>
</tr>
<tr>
<td>Pack-years smoking</td>
<td></td>
<td>(&lt;0.01)</td>
</tr>
<tr>
<td>Never smoker</td>
<td>326 (68)</td>
<td>Ref.</td>
</tr>
<tr>
<td>Lower tertile (&gt;0 – 6.25)</td>
<td>66 (14)</td>
<td>(-0.10 ± 0.038^{**})</td>
</tr>
<tr>
<td>Middle tertile (&gt;6.25 – 16)</td>
<td>43 (9)</td>
<td>(-0.14 ± 0.046^{**})</td>
</tr>
<tr>
<td>Higher tertile (&gt;16)</td>
<td>42 (9)</td>
<td>0.009 ± 0.048</td>
</tr>
<tr>
<td>(p_{trend})</td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>Total physical activity</td>
<td></td>
<td>0.12</td>
</tr>
<tr>
<td>Low</td>
<td>42 (9)</td>
<td>Ref.</td>
</tr>
<tr>
<td>Moderate</td>
<td>225 (47)</td>
<td>0.050 ± 0.048</td>
</tr>
<tr>
<td>High</td>
<td>210 (44)</td>
<td>0.089 ± 0.048</td>
</tr>
<tr>
<td>(p_{trend})</td>
<td></td>
<td>(0.04)</td>
</tr>
<tr>
<td>Vigorous physical activity</td>
<td></td>
<td>(&lt;0.01)</td>
</tr>
<tr>
<td>(MET-mins)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1 (0)</td>
<td>147 (31)</td>
<td>Ref.</td>
</tr>
<tr>
<td>Q2 (&gt;0- &lt;840)</td>
<td>91 (19)</td>
<td>0.054 ± 0.038</td>
</tr>
<tr>
<td>Q3 (840- &lt;1920)</td>
<td>134 (28)</td>
<td>0.013 ± 0.034</td>
</tr>
<tr>
<td>Q4 (≥1920)</td>
<td>103 (22)</td>
<td>(0.13 ± 0.037^{**})</td>
</tr>
<tr>
<td>(p_{trend})</td>
<td></td>
<td>(&lt;0.01)</td>
</tr>
</tbody>
</table>

Abbreviations: SE=Standard error; Ref.=Referent

\(^a\)The age adjusted model adjusted for age and included a random-effects parameter representing study center.
The fully adjusted models adjusted for age, sex, and lipid ratio while also adjusting for the other lifestyle exposures of interest. For example, the model with alcohol consumption adjusts for age, sex, lipid ratio, smoking status, and total physical activity. These models also included a random-effects parameter representing study center.

Overall p-value computed by F-statistic.

* p<0.05 in comparison to referent computed by t-statistic

** p<0.01 in comparison to referent computed by t-statistic
References


Chapter 5

Gene-Environment Interaction: Alcohol Consumption and CYP2E1*5B

5.1 Introduction

Alcohol is an established risk factor for cancer and has been classified as a group 1 carcinogen by the International Agency for Research on Cancer (1). The health effects of alcohol are related to the byproducts generated during alcohol metabolism (2). Alcohol is predominantly metabolized by the enzyme alcohol dehydrogenase (ADH); however, cytochrome P450 2E1 (CYP2E1)-mediated alcohol metabolism becomes increasingly more important as higher levels of alcohol are consumed (3). Polymorphic differences in alcohol metabolizing enzymes can affect the rate of alcohol metabolism and the carcinogenic dose for the same volume of alcohol consumed by different individuals. Several polymorphisms have been identified for CYP2E1 of these CYP2E1*5B (rs3813867) is the most commonly studied. CYP2E1*5B has been linked to increased cancer risk and in some studies has modified the effect of alcohol consumption on cancer (4). While some ambiguity still exists, the CYP2E1*5B polymorphism appears to affect gene expression and enzyme catalytic activity (5-7). This may lead to greater CYP2E1-mediated alcohol metabolism under inducing conditions, such as during periods of high alcohol intake and an increased concomitant production of ROS in individuals with the variant form compared to individuals with the normal, wild-type form of the gene. To our knowledge there are currently no published studies on the effect of the CYP2E1*5B polymorphism on telomere length or on the potential modifying effect of CYP2E1*5B on the relation between alcohol and telomere length. This chapter presents an exploratory analysis of a gene-environment interaction between alcohol consumption and CYP2E1*5B polymorphism.
5.2 Methods

5.2.1 CYP2E1 Genotyping

CYP2E1 genotyping was completed by the Queen’s Laboratory for Molecular Pathology using the TaqMan® Drug Metabolism Genotyping Assay (Assay ID: C_2431875_10, Applied Biosystems). This PCR based assay amplifies and quantifies targeted DNA sequences using sequence-specific TaqMan® probes. When the probe hybridizes to the target DNA sequence a fluorescence signal is detected permitting the determination of different genotypes of CYP2E1. To assess the validity of genotyping methods, PCR products were sequenced to confirm specificity of the reaction. Additionally, repeat assays were conducted for 10% of samples, at random, and concordance between assays was evaluated. The between-assay agreement was 100%. Departures from Hardy-Weinberg equilibrium may indicate systematic genotyping errors. Hardy-Weinberg equilibrium was confirmed by comparing the consistency of the observed genotype frequencies against the expected under equilibrium using chi-squared goodness-of-fit and Fisher’s exact test.

5.2.2 Analysis Strategy for Assessment of Interaction

To investigate the gene-environment interaction a product term was included in a multiple linear regression model that was adjusted for age and that included a random effects parameter representing study center. Due to the small sample size in the variant category other covariates were not considered or adjusted for in this exploratory analysis. The distribution of CYP2E1 genotype was such that there were no individuals with the homozygous recessive (*5B/*5B) genotype in the study population. Since the *5B/*5B category had a zero cell count it was collapsed with heterozygous recessive (*1A/*5B) genotype for the purposes of this analysis.
5.3 Results

Table 5-1 presents the distribution of alcohol consumption dichotomized at 7 drinks per week and CYP2E1 genotype in the study population. Age-adjusted regression coefficients for telomere length are also presented. Approximately 10% (n=50) of the study population had the *1A/*5B genotype, while 90% (n=427) of the population was homozygous wild-type (*1A/*1A). Alcohol consumption and CYP2E1 were not associated with telomere length.

Results of the interaction analysis are presented in Table 5-2. Qualitatively, the direction of the regression coefficient was different across categories of CYP2E1 genotype and the magnitude of the effect of alcohol consumption on rLTL was stronger in the variant genotype stratum compared to the normal, wild-type genotype stratum. However, only a small proportion of study participants were carriers of the CYP2E1*5B polymorphism and an even smaller number of participants were carriers of the polymorphism and also consumed more than seven alcoholic beverages per week. As a result, the presented regression coefficient and p-value are interpreted with caution. The direction and magnitude of the mean change in rLTL for those drinking more than 7 drinks/week compared to those drinking fewer than 7 drinks/week suggest a protective association between alcohol consumption and rLTL in participants with the CYP2E1*5B polymorphism. This association was not significant (p=0.21). Among participants with the normal, wild-type genotype drinking more than 7 drinks/week was associated with on average 0.057 relative T/S ratio unit shorter rLTL compared to those consuming 7 or fewer alcoholic beverages/week; however, this association was also not significant (p=0.11).

The distribution of the CYP2E1*5B polymorphism was in Hardy-Weinberg equilibrium (Fisher’s exact p=0.89, data not shown).
5.4 Discussion

The CYP2E1*5B polymorphism was not associated with rLTL and a significant interaction with alcohol consumption was not observed. However, this analysis was underpowered to detect a significant association should one exist. Only 10% of study participants were carriers of the heterozygous variant genotype and the proportion of participants with the variant genotype that also consumed more than seven drinks per week was even smaller (2%). The direction and magnitude of the regression coefficients in stratified analysis suggest a stronger positive relationship between alcohol and rLTL among those with the variant genotype. Studies that considered effect modification by CYP2E1*5B on the relationship between alcohol and cancer have produced conflicting results (4). These inconsistencies are likely due to the varying prevalence among ethnicities; the prevalence of CYP2E1*5B is 2-5% in Caucasian populations (7,8). However, studies that reported on the relation between CYP2E1*5B and cancer stratified by alcohol consumption typically reported a stronger association between CYP2E1*5B and cancer among those who consumed higher levels of alcohol (4).

Alcohol metabolism via the CYP2E1-pathway can be induced 10 to 20-fold with greater blood alcohol levels and this induction may occur at an approximate dose of three drinks per day (9, 10). Alcohol metabolism by CYP2E1 results in a significant generation of reactive oxygen species, which can lead to increased lipid peroxidation and DNA damage (11). Additionally CYP2E1 has been shown to activate a variety of pro-carcinogens (12). While the relative contribution of these mechanisms to telomere attrition and cancer is unknown, current evidence supports the hypothesis that the generation of ROS may contribute to telomere shortening (13, 14). Functional polymorphic differences in CYP2E1 can cause changes in enzyme activity and enzyme sensitivity (6, 15) and thus may affect the rate of ROS production. The CYP2E1*5B
variant is associated with higher transcriptional activity, higher protein levels, and higher enzyme activity compared to the normal genotype (15). However, the prevalence of CYP2E1*5B can vary substantially between ethnicities (8, 16) and this may explain the inconsistencies that exist in the literature. We were not able to make substantive inferences about the effect of the CYP2E1*5B polymorphism on telomere length or any effect modification on the relationship between alcohol consumption and telomere length due to a small number of participants with this variant in our study population.
5.5 Figures and Tables

Figure 5-1. Alcohol metabolism via the ADH-ALDH pathway. Alcohol is oxidized by alcohol dehydrogenase (ADH) to the carcinogenic intermediate acetaldehyde. Acetaldehyde is then further oxidized to acetate, by aldehyde dehydrogenase (ALDH), which is eventually converted to acetyl CoA. Depending on nutritional status and hormone regulation acetyl CoA is converted to any of carbon dioxide, ketone bodies, fatty acids and/or cholesterol.
Figure 5-2. CYP2E1 metabolism of alcohol. CYP2E1 enzyme has the highest activity for oxidizing alcohol to acetaldehyde but the alcohol concentration needs to be much higher for alcohol to be preferentially metabolized by CYP2E1 over ADH. However, the alcohol levels required to active the CYP2E1 pathway are within the range of alcohol concentrations seen in social drinking. The relevance of CYP2E1 in alcohol metabolism increases as blood concentration of alcohol increases. Variation in CYP2E1 genotype may affect the rate of alcohol metabolism, the generation of oxidative stress and may therefore modify the relationship between alcohol consumption and leukocyte telomere length.
Table 5-1. Age-adjusted regression coefficients and associated standard errors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N (%)</th>
<th>Coefficient ± SE</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol (drinks/wk)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 7</td>
<td>370 (76)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>&gt; 7</td>
<td>107 (22)</td>
<td>-0.038 ± 0.031</td>
<td></td>
</tr>
<tr>
<td>CYP2E1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1A/ *1A</td>
<td>427 (90)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>*1A/ *5B</td>
<td>50 (10)</td>
<td>-0.0052 ± 0.043</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Ref. = referent; SE=standard error
Table 5-2. Exploratory analysis of effect modification by the CYP2E1*5B polymorphism on the relationship between alcohol consumption and rLTL. Alcohol consumption was dichotomized at 7 drinks/week for this analysis. *1A/*5B is the heterozygous variant and *1A/*1A is the homozygous wild-type.

<table>
<thead>
<tr>
<th>Alcohol consumption</th>
<th>CYP2E1</th>
<th>Alcohol consumption</th>
<th>CYP2E1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>*1A/ *5B</td>
<td>(N=50)</td>
<td>*1A/ *1A</td>
</tr>
<tr>
<td>≤ 7 drinks/ week (N=370)</td>
<td>Ref.</td>
<td>-</td>
<td>N=99</td>
</tr>
<tr>
<td>&gt; 7 drinks/ week (N=107)</td>
<td>0.14 ± 0.11</td>
<td>0.21</td>
<td>N=8</td>
</tr>
</tbody>
</table>

Abbreviations: Coeff.=Coefficient; Ref. = referent; SE=standard error
p-value for interaction = 0.10
References


Chapter 6

General Discussion and Conclusions

The broader goal of this research was to investigate lifestyle determinants of telomere length. Specifically, the purpose of this thesis was to determine the relationship between telomere length and alcohol consumption, smoking and physical activity. This study was nested within a larger cross-sectional study that recruited healthy volunteers from Ottawa, Ontario, Kingston, Ontario, and Halifax, Nova Scotia. The following sections highlight the main findings of this thesis work and discuss limitations, strengths and generalizability. A brief discussion of future work and a summary of the main conclusions are also included.

6.1 Summary of Main Findings and Comparison to the Literature

6.1.1 Alcohol Consumption and Interaction with CYP2E1

Alcohol consumption was not associated with relative leukocyte telomere length in this population of healthy volunteers aged 20 to 50 years. Two other studies have considered the relationship and both reported an inverse association between alcohol consumption and telomere length; a cross-sectional analysis in a group of Italian men diagnosed with alcohol abuse (1) and a retrospective longitudinal study in older Helsinki businessmen (2). Both studies recruited solely male participants limiting generalizability. Pavenello et al. (1) reported nearly halved rLTL in those with alcohol abuse compared to social drinkers but did not observe an association between rLTL and alcohol consumption when evaluating social drinkers separately. Strandberg et al. (2) reported a decreasing linear trend in rLTL at old age with increasing alcohol consumption at midlife. It is possible that this study measured alcohol intake at a more biologically relevant time. Detailed alcohol intake was measured at midlife, which is likely representative of peak alcohol
consumption during healthy life and may be a better approximation of a cumulative lifetime dose. Considering that telomere length shortens with age it may be more appropriate to measure cumulative lifetime dose of exposure. Strandberg et al. (2) did not observe an association in cross-sectional analysis at old age, likely due to decreasing alcohol consumption with intervening diseases at older age.

Alcohol is preferentially metabolized to other nutrients, consequently in the presence of high blood alcohol levels the consumption of normal nutrients is significantly decreased (3). Because persons with alcohol abuse may consume considerably more alcohol than social drinkers, and alcohol is preferentially metabolized to other nutrients, the nutrient profiles of these two populations may be considerably different. This may explain the observed association in those with alcohol abuse but not in more moderate, social drinkers. Evidence suggests that healthy behaviours cluster in protected drinkers, those groups that show health benefits related to alcohol consumption, and this cannot be adjusted away using statistical methods (4-7). Further, plasma concentrations of antioxidants such as vitamin E, vitamin C and selenium are lower in those with alcohol abuse than in those who consume moderate amounts of alcohol, whereas markers of oxidative stress such as malondialdehyde and autoantibodies directed at malondialdehyde are higher (8). While our results suggest that moderate alcohol consumption is not associated with relative leukocyte telomere length evidence from previous studies suggests an association with heavy drinking.

Effect modification by CYP2E1 on the relation between alcohol and telomere length was considered. However, this study population had a very small number of participants with the CYP2E1*5B variant and no substantive conclusions could be made.
6.1.2 Cigarette Smoking

Both current smoking and cigarette pack-years were inversely associated with rLTL. The existing literature corroborates an association between leukocyte telomere length and smoking. Six studies have reported on smoking as a primary exposure of interests, three of cross-sectional design, one case-control and two longitudinal studies. Of the three cross-sectional studies, two (9, 10) reported a negative correlation and one (11) reported an unadjusted negative correlation between pack-years smoking and rLTL. The case-control study (12) reported higher odds of having rLTL below the median in ever versus never smokers. One retrospective study (13) reported shorter rLTL in old age in relation to smoking at healthy midlife and the other longitudinal study (14) reported an approximately tripled telomere attrition rate per year in active smokers compared to never smokers. This study (14) also reported that while baseline rLTL was shorter in former smokers, their annual telomere attrition rate was comparable to that of participants who had never smoked.

Cigarette smoking is an important risk factor for many age-related diseases and is associated with low-grade, systemic oxidative stress and inflammation (15-19). Additionally, cigarette smoke can induce single-strand break damage in the DNA of human cells (20). The results of this thesis work indicate shorter telomere length in smokers compared to non-smokers suggesting that the rate of telomere shortening may vary between individuals with different degrees of exposure to cigarette smoke induced oxidative stress.
6.1.3 Physical Activity

A linear trend in increasing telomere length was observed with increasing total physical activity; however this trend was not significant at p=0.06. Greater vigorous physical activity was associated with longer telomeres. An association with physical activity is supported by several studies (21-26) but not by others (27, 28) who reported a null association with telomere length. A lack of consensus exists in the literature regarding the intensity of physical activity that confers the greatest benefit on rLTL. The results of this thesis work are consistent with reports of a linear increase with greater vigorous physical activity (21, 22). Physical activity is associated with many health benefits and may favourably alter the oxidant-antioxidant balance. Exposure to a 16-week physical activity program in healthy young men and women increased endogenous antioxidant activity, specifically glutathione peroxidase and glutathione reductase, resistance to LDL oxidation and consequently decreased oxidized LDL concentration (29). The Harvard Alumni Study (30) investigated the relationship between exercise and longevity and reported a graded inverse relationship between mortality and total physical activity. Additionally, vigorous physical activity was associated with longevity but non-vigorous physical activity was not (30). The mechanism behind these relationships is not completely understood; however, one study has suggested that favourable changes in high-density lipoprotein (HDL) cholesterol and triglyceride levels occur at a threshold intensity of 5 to 6 METs (31). Others have argued that the benefits of vigorous physical activity may be a result of improved cardiorespiratory conditioning; thus, sustained and dynamic physical activity may have greater benefit than intermittent or less intense exercise (32, 33). Ultimately, the relationship between physical activity and rLTL is likely to be U-shaped as the beneficial effects of exercise taper off with prolonged physical exertion (29, 34).
Together, these results suggest that cigarette smoking and physical activity may share metabolic pathways that affect telomere length, likely mediated by oxidative stress and inflammation.

6.2 Sensitivity Analysis and the Robustness of Violations of Normality Assumption

The t- and F-tests are fairly robust to violations of the normality assumption in larger samples; the type-I error rate or probability of rejecting the null hypothesis when it is true is not seriously affected. A Q-Q plot of the residual distribution is presented in Appendix D. The violation to normality is moderate and it is likely that with this large sample size the distribution of the differences between group means converges to a normal distribution (35). A sensitivity analysis with natural log-transformed telomere length demonstrated qualitatively the same results as those obtained from the models with untransformed telomere length.

6.3 Limitations and Strengths

6.3.1 Limitations of Study Design

The population for this study was comprised of healthy volunteers and may not be representative of the general population. However, the goal of this study was to assess a biological relationship between lifestyle factors and telomere length and not to describe the prevalence of alcohol consumption, smoking, physical activity or telomere length in general. Therefore, the use of a volunteer population does not detract from a contribution in the understanding of the underlying relationship of interest. However, the recruited sample represents a very healthy population and as a result we observed a narrower range of variation in the lifestyle exposures under consideration.
A general disadvantage of cross-sectional studies is the inability to confirm cause-and-effect relationships as exposure and outcome are measured at the same time and chronological sequence cannot be unequivocally established. It is unlikely that telomere length can influences a person’s drinking patterns, smoking habits, or physical activity levels and we can eliminate, with reasonable certainty, the possibility of reverse causality. However, other aspects of the cross-sectional study design limit causal inference, including the measurement of lifestyle factors and telomere length at a single point in time. A one-time measure of smoking and in particular alcohol consumption and physical activity may not capture a cumulative lifetime exposure measure which may be a more relevant measure for the study of their effect on telomere length. A single measure may also increase the risk of misclassification of study subjects. Additionally, telomere length was measured in leukocyte DNA extracted from a single blood draw. Both biological and technical factors can introduce variation in the measurement of telomere length at a single point in time. This inherent imprecision in a single measurement may influence the observed estimates of association (36). Using an average of multiple measurements over a short period of time may provide a more precise measure of rLTL; however, this is often impractical in large epidemiological studies. Kim et al. (36) demonstrated a good short-term reliability of telomere length over a period of nine-months and also reported a lack of seasonal variation in telomere length measurement.

Unmeasured and residual confounding are a common problem in establishing causal inference in observational studies. Many aspects of telomere length biology are still unknown and measurement error in confounders can result in residual confounding. Depending on the amount and direction of bias effect estimates, there is the potential that the effect sizes, of the magnitude reported here, are a result of unmeasured and residual confounding alone.
6.3.2 Limitations of Exposure Measurement

Alcohol consumption, smoking and physical activity were assessed as covariates in the larger study, therefore, not all dimensions of these lifestyle exposures are presented. Additionally, all were assessed over the past month. A shorter recall period would not have been appropriate, as it would likely not have accurately represented the respondents’ lifetime exposure. This is particularly true for alcohol consumption and physical activity patterns, which can be highly variable. A longer reference period of one or two years may have allowed for more appropriate estimates of cumulative lifetime exposure in this population. Alternatively, the questionnaire could have included questions about alcohol consumption, smoking and physical activity habits during different periods of life such as adolescence, young adulthood, middle and late adulthood.

Longitudinal panel studies have shown high correlation between alcohol consumption measurements for adult samples five years apart or less (37, 38). However, heavy drinkers may have less stable drinking patterns than abstainers or moderate drinkers (37). Use of the one-month reference period may have increased the likelihood of misclassifying infrequent drinkers as abstainers. However, all analysis compared to the ‘low’ alcohol consumption category and not “abstainer” reducing the impact of this misclassification on the presented results. Studies of the stability of general physical activity patterns across the lifespan suggest that physical activity levels are somewhat stable from childhood through middle and late adulthood (39, 40). Additionally, children and adolescents who participate in sports are more likely to remain physically active in adulthood (41). The validity of adult self-reported smoking in observational studies has also been demonstrated (42, 43). However, current smoking was defined as currently smoking at least one cigarette per day and former smoking was defined as having smoked at least
one cigarette per day for a period of at least six months. Thus, it is likely that any occasional smokers were classified as never smokers.

Alcohol intake, physical activity and smoking were self-reported rendering the measures obtained susceptible to misclassification and social desirability bias (44-46). While some non-differential misclassification of exposure is likely, currently self-reported measures are considered the best method of retrospectively measuring alcohol consumption (47-50). Alternatives such as direct observation are impractical and measure very recent drinking (48). Self-reports of alcohol consumption conducted in clinical or research settings, where subjects are given assurances of confidentiality, are reliable tools for measuring alcohol consumption at the individual level (48, 49). Poikolainen et al. (50) estimated that the quantity-frequency questionnaire method of obtaining self-reported alcohol intake is closely correlated with daily diary intake (r > 0.90) and commonly used laboratory markers (r=0.41 - 0.67).

Quantity-frequency methods of assessing alcohol consumption are often criticized for not measuring atypical heavy drinking and variable patterns of drinking (48). Studies that assess heavy drinking episodes in normal drinkers and that include beverage-specific quantity-frequency questions generate higher estimates of alcohol intake than simpler quantity-frequency questionnaires (47-49). The average number of days that participants engaged in binge drinking or the average amount of alcohol consumed on binge-drinking occasions was not assessed in this study. Stockwell et al. (51) sate that while most people drink in excess of the guidelines for acute risk occasionally, a relatively small proportion report doing so regularly enough to cause long term health effects, such as cancer, and it is typically young drinkers that drink above the acute risk guidelines.
The repeatability of the IPAQ instrument was assessed by Craigs et al. (53) who compared assigned physical activity levels based on two completions of the IPAQ questionnaire no more than ten days apart. Correlation between repeat assessments was 80%. The IPAQ was also compared to CSA accelerometer data obtained over a seven day period. The IPAQ and accelerometer assessments were 30% correlated, comparable to other self-report validation studies (53). While the IPAQ short form may in some instances overestimate overall physical activity and classify less active individuals as highly active (52) it has been shown to be equal to other surveys of self-reported physical activity (53). Additionally, the reliability and validity of the vigorous activity metric is greater than that of total physical activity (54). As with alcohol our measure of physical activity may not be truly representative of a lifetime cumulative exposure and our results probably reflect the cross-sectional as opposed to long-term effects of exercise on telomere length.

The widespread implementation of public health policies to reduce smoking may reinforce the perception of smoking as a socially undesirable behaviour and increase bias in self-reported smoking status (55). Prevalence estimates of self-reported cigarette smoking are slightly lower compared to those of urinary cotinine concentration, a major metabolite of nicotine (55). However, sensitivity estimates demonstrate that 90% of adults who were classified as smokers based on urine cotinine concentration also reported current smoking (55). This demonstrates that accurate estimates of the prevalence of cigarette smoking among study participants can be derived from self-reported smoking status.

As most individuals do not know their telomere length and this outcome was not advertised on study posters it is likely that any misclassification is non-differential. This type of
misclassification is most likely to increase the error term and make it harder to detect differences (56).

6.3.3 Limitations of Telomere Length Measurement

Telomere measurement by quantitative PCR is simple, fast, high-throughput and easily scalable for large studies; however, it quantifies the average telomere length per sample and cannot quantify individual telomeres (57). Additionally this method produces relative telomere length measures and cannot provide absolute telomere length values in kilobases (57) or the proportion of short telomeres. Therefore, qPCR is only able to detect large differences in telomere length as it provides an estimate of the mean telomere length of the whole cell population (57). Methods such as Q-FISH are able to quantify the number of cells with critically short telomeres, which may be a more appropriate measure for this question. However, this method requires well-prepared metaphase cells which were not available for this study, significantly more resources, is more labour intensive and not well suited for high throughput analysis.

The MMQPCR method has larger laboratory measurement error than the Southern blot method of TRFs. However, precision alone might not justify using the Southern blot method as it requires large amounts of DNA, 3 µg compared to 20 ng required for qPCR (58), cannot be performed on DNA that has even modest degradation (59), and is appreciably more labour intensive and costly. Furthermore, the restriction enzymes used to generate TRFs target sub-telomeric regions with polymorphic sites, which may vary across the population while qPCR measures only the canonical region of telomeres (60, 61).

An important assumption of the MMQPCR assay is that measures taken on DNA of different quality are comparable. Further, standardization of the T/S ratio for each sample to a reference sample does not ensure comparability between different research groups. The use of a
synthesized oligonucleotide of known length as the reference would have allowed for the calculation of absolute telomere length and facilitated comparability to other studies that also estimated absolute telomere length values.

Telomere length measured by qPCR has been shown to be strongly correlated with the traditional Southern blot method for measuring TRF lengths ($R^2=0.844$ (62); $r>0.9$ (61); $r=0.88$ (64)). Kim et al. (36) reported an intraclass correlation coefficient of 0.64 for telomere length measurements over seven visits spanning a nine month period; demonstrating good stability of telomere length measurement over a period of months. These results emphasize that a single blood draw serves as a reliable measure of telomere length.

6.3.4 Strengths

This study explores the relationship between prevalent and modifiable lifestyle exposures and a biomarker of early effect, which is suspected to play an important role in carcinogenesis. Determinants of telomere length are poorly understood and much of the existing literature is inconsistent. This thesis examined the relationship between telomere length and three lifestyle factors deemed to be important in cancer etiology: alcohol consumption, cigarette smoking, and physical activity and thus contributes knowledge to a relatively new area of research with limited existing knowledge. Issues of temporality, typical of most cross-sectional analysis, do not apply here unless we invoke the unlikely possibility that individuals with shorter telomeres are prone to consume more alcohol, smoke more, or exercise less throughout life. The narrower age range of 20 to 50 years is a strength of this study as telomere length exhibits instability and steeper loss in the very young or very old, while midlife is marked by a gradual loss in telomere length. This allowed us to model telomere length during a period of life where telomere loss is gradual, may be better represented by a linear model, and where age may have a smaller influence.
6.4 Generalizability

This study population was comprised of healthy men and women aged 20 to 50 years. Results are not generalizable to populations outside this age range and these biologic relationships are most likely to hold in healthy populations with a similar age and exposure range. It is possible that these relationships may be affected by unknown genetic factors that may differ across populations. As this study population was primarily Caucasian results may be different in populations with varying ethnicities.

6.5 Clinical Significance of Telomere Length

What constitutes a clinically significant telomere length is currently not known. The extent of the malleability of telomere length is also still unclear. There is no established standard for a critically short telomere length and whether changes in behavioural and environmental factors can improve telomere length in a clinically meaningful way that results in a reduced risk of disease remains to be determined. A comparison of regression coefficients obtained in this study for the lifestyle exposures with that observed for age provides some interpretation of the clinical significance of these lifestyle factors in relation to telomere length. The coefficient for vigorous physical activity suggests a 20 year decrease in biological age for those participating in 60 minutes of vigorous physical activity 4 times a week and the coefficient for current smoking suggests a 16 year increase in ‘biological age’ for current smokers compared to never smokers.

While telomeres will shorten over a period of decades it has been hypothesized that sustained shorter-term increases in telomere length may alter the long-term trajectory of telomere attrition in such a way that decreases risk of age-related diseases (65). This thought is based on observations of increased lifespan in animal models with upregulated telomerase activity (66) and improvements in health that coincided with a modest increase in telomerase following...
behavioural interventions (67, 68). With telomere length dichotomized at the median, those with longer telomere length lived an average of five years longer than those with shorter telomeres (69). This may be due to a delay in senescence and associated pathology with longer telomeres.

6.6 Future Directions

Future studies should focus on the use of prospective study designs that measure cumulative lifetime exposure and take multiple telomere length measures over a period of years. Unmeasured confounding and statistical power are issues in many studies and the careful consideration of statistical power and the careful measurement of potential confounders are additional study design considerations that are important for future work in this area.

Evidence supporting telomere length as a biomarker of increased cancer risk is mainly derived from retrospective case-control studies, with smoking and inflammation-related cancers showing the strongest and most consistent results. Prospective studies conducted on the relationship are more inconsistent and suggest that instead of being a general biomarker of increased risk, leukocyte telomere length may be associated with certain cancer types or may be prognostic and indicative of survival with cancer (62). Reverse causation bias may be present in retrospective studies that collected samples for cases after cancer diagnosis as it is possible that the presence of the tumour, the psychological impact of the diagnosis and living with the disease, and DNA damage resulting from treatment had an impact on telomere length at the time of sample collection. Longitudinal studies of cancer risk that collect multiple specimens prior to cancer development are needed to address this issue. Use of consistent tissue and cell type, consistent extraction and storage of DNA, consistent use of telomere length assay methods, detailed descriptions of and consistent analytic design, as well as the careful measurement of important predictors and interacting factors is crucial for meaningful results in future studies.
Large, population-based longitudinal studies are needed to define a typical telomere length range for an individual’s age and intervention studies can help identify and test effective components for change, to what extent telomere length can be modified, and whether these changes translate to a meaningful change in well-being and reduced risk of disease.

6.7 Conclusions

Smoking and physical activity are modifiable factors that have an effect on telomere length and consequently may influence the ageing process and the progression to ageing-related disease. Leukocyte telomere length may mediate the relationship between lifestyle and the risk of cancer and other age-related diseases. This message can be used by clinicians and public health professionals to promote smoking cessation and regular exercise.

The scientific evidence relating lifestyle factors such as alcohol consumption, smoking and physical activity to cancer is pretty convincing. Despite this, the percent of Canadians exceeding low risk drinking guidelines has steadily increased, smoking continues to be a major contributor to morbidity and mortality, and most Canadians do not meet minimum physical activity guidelines. Elucidating the mechanisms that underlie the association between these prevalent risk factors for cancer and an intermediate in this pathway may allow for intervention at a time when it is still possible to prevent disease onset. Additionally, in research telomere length can be used as a health outcome that can be assessed on a shorter temporal scale.
References


17. Alberg A. The influence of cigarette smoking on circulating concentrations of antioxidants micronutrients. Toxicology. 2002; 180:121-137.


Appendix A

Health Sciences Research Ethics Board Approval

QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD-DELEGATED REVIEW
September 27, 2012

Miss Lidija Latifovic
Department of Community Health and Epidemiology
Carruthers Hall
Queen’s University
Kingston ON K7L 3N6

Dear Dr. Latifovic:

Romeo File # 6007409 EPID-397-12 Study Title: The influence of alcohol consumption and lifestyle factors on telomere length in peripheral blood leukocytes.
Co-Investigators: Dr. Will D King, Dr. Thomas E Massey

I am writing to acknowledge receipt of your recent ethics submission. We have examined the protocol and consent form for your project (as stated above) and consider it to be ethically acceptable. This approval is valid for one year from the date of the Chair’s signature below. This approval will be reported to the Research Ethics Board. Please attend carefully to the following listing of ethics requirements you must fulfill over the course of your study:

Reporting of Amendments: If there are any changes to your study (e.g. consent, protocol, study procedures, etc.), you must submit an amendment to the Research Ethics Board for approval. Please use event form: HSREB Multi-Use Amendment/Full Board Renewal Form associated with your post review file # 6007409 in your Researcher Portal (https://eservices.queensu.ca/romeo_researcher/)

Reporting of Serious Adverse Events: Any unexpected serious adverse event occurring locally must be reported within 2 working days or earlier if required by the study sponsor. All other serious adverse events must be reported within 15 days after becoming aware of the information. Serious Adverse Event forms are located with your post-review file 6007409 in your Researcher Portal (https://eservices.queensu.ca/romeo_researcher/)

Reporting of Complaints: Any complaints made by participants or persons acting on behalf of participants must be reported to the Research Ethics Board within 7 days of becoming aware of the complaint. Note: All documents supplied to participants must have the contact information for the Research Ethics Board.

Annual Renewal: Prior to the expiration of your approval (which is one year from the date of the Chair’s signature below), you will be reminded to submit your renewal form along with any new changes or amendments you wish to make to your study. If there have been no major changes to your protocol, your approval may be renewed for another year.

Yours sincerely,

Chair, Research Ethics Board
September 27, 2012

Investigators please note that if your trial is registered by the sponsor, you must take responsibility to ensure that the registration information is accurate and complete

126
Appendix B

Monochrome Multiplex qPCR Protocol

Triplicate samples of 20 ng of DNA were amplified using telomere primers and single copy gene human beta globin primers manufactured by ID Technologies. Human beta globin primers were modified with a 5’ GC clamp that shifted melting temperature. This modification ensured that beta globin DNA amplification and signal acquisition occurred at a higher temperature after telomere product had melted to baseline and was undetectable.

Primer Sequences

Telomere primer pair:

\[
\begin{align*}
telg & : \text{ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT} \\
telc & : \text{TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCTATCCCTAAACA}
\end{align*}
\]

Single copy gene (human beta globin) primer pair:

\[
\begin{align*}
hbgd & : \text{GCCCGGCCGGCGCCGCCTCGCCGCGGGGgaggagaagttgctggtt} \\
hbgu & : \text{CGGCGGCAGGCGGCGGCGGGTGGGCGGtgcctacggtcaccttg}
\end{align*}
\]

The lower case letters indicate the 5’ GC clamp added to the primer to shift its melting temperature. This allowed for multiplexing of the two reactions.
Table B-1. Protocol for monochrome multiplex quantitative PCR using human betaglobin as the single copy gene. Each well had a final volume of 25 µl, which included 15 µl master mix (primers and polymerase included) and 10 µl sample DNA.

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Master Mix</strong></td>
<td></td>
</tr>
<tr>
<td>SYBR Green I (Life Technologies)</td>
<td>0.75 x</td>
</tr>
<tr>
<td>Tris-HCl, pH range 8.3 - 8.5 (Sigma) a</td>
<td>10 mM</td>
</tr>
<tr>
<td>KCl (Sigma)</td>
<td>50 mM</td>
</tr>
<tr>
<td>MgCl₂ (Sigma)</td>
<td>3 mM</td>
</tr>
<tr>
<td>dNTP (each) (Life Technologies)</td>
<td>0.2 mM</td>
</tr>
<tr>
<td>DTT (Sigma)</td>
<td>1 mM</td>
</tr>
<tr>
<td>Betaine (Sigma)</td>
<td>1 M</td>
</tr>
<tr>
<td><strong>Primers</strong></td>
<td></td>
</tr>
<tr>
<td>telg (IDT)</td>
<td>100 nM</td>
</tr>
<tr>
<td>telc (IDT)</td>
<td>900 nM</td>
</tr>
<tr>
<td>hbgd (IDT)</td>
<td>500 nM</td>
</tr>
<tr>
<td>hbgu (IDT)</td>
<td>500 nM</td>
</tr>
<tr>
<td><strong>Polymerase</strong></td>
<td></td>
</tr>
<tr>
<td>AmpliTaq Gold (Life Technologies)</td>
<td>0.025 U/µl</td>
</tr>
</tbody>
</table>

Molecular grade water (VWR)

a pH was estimated using standard full-range pH indicator paper (Sigma)
Table B-2. Thermal cycling profile for the MMQPCR reaction using human betaglobin as the single copy gene. The 74°C signal acquisition provides the Cq values for the amplification of telomere template (single copy gene signal is still at baseline) and the 85°C signal acquisition provides the Cq values for the amplification of the single copy gene (human betaglobin) template. At 85°C, there is no signal from the telomere product because it is fully melted.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Temperature</th>
<th>Time</th>
<th>Number of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Activating polymerase</td>
<td>95°C</td>
<td>15 min</td>
<td>1</td>
</tr>
<tr>
<td>2: Denaturation</td>
<td>98°C</td>
<td>2 sec</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>49°C</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td>3: Denaturation</td>
<td>98°C</td>
<td>2 sec</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>59°C</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>74°C</td>
<td>15 sec</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Signal Acquisition</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>84°C</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>85°C</td>
<td>15 sec</td>
<td></td>
</tr>
<tr>
<td>4: Melt curve</td>
<td>72-95°C</td>
<td>0.5°C/ 30 sec</td>
<td>1</td>
</tr>
</tbody>
</table>


Gel Recipe and Preparation

Recipe for 10X TBE Buffer:
- Add the following to 800 mL H₂O:
  - 108 g Tris Base
  - 5.5 g boric acid
  - 9.3 g EDTA (or 7.5 g of EDTA, disodium dehydrated dehydrate (salt))

Recipe for Agarose Gel (3%):
- Add 1.8 g agarose to 50 mL of 1X TBE Buffer in a clean flask
- Adjust volume to 60 mL with additional 1X TBE buffer or H₂O
- Microwave at 30 second intervals until agarose is completely dissolved (takes about 2.5 minutes)
  - swirl to mix
  - cover in microwave with a small beaker
- Let gel cool for 1-2 minutes
- Add 1.0 µl ethidium bromide to agarose – TBE mixture
  - swirl to mix
- Pour mixture into casting tray
  - prior to pouring gel ensure tray is level and insert rubber stoppers
  - after pouring mixture into tray add comb
- Allow 15-20 minutes for gel to harden
- Remove comb and rubber stopper
- Cover with 1X TBE

Gel Electrophoresis:
- Place 1 µl of loading dye on paraffin (or 3 µl of 1 in 3 dilution of loading dye), 1 dot of dye for each sample to be loaded
- Aliquot 8 µl of sample and mix with dye on paraffin by pipetting up and down several times
- Add sample and dye mixture to wells
- Load 6 µl of DNA ladder to last well
- Set up casting tray so that wells are at the black end as current will run from black to red
- On power supply plug black wire into black and red wire into red
- Set voltage to 85 V for 1 to 1.5 hours
Appendix C

Study Questionnaire for Smoking, Alcohol Consumption, and Physical Activity

SMOKING AND ALCOHOL CONSUMPTION

15. During the past month have you smoked at least 1 cigarette a day?
   □ No
   □ Yes → on average, how many cigarettes did you usually smoke in one day?
            ___ cigarettes per day

16. Have you ever smoked at least 1 cigarette a day for 6 months or more?
   □ No
   □ Yes → a. How old were you when you began to smoke? ___ years old
   □ No → b. When you smoked the most, how many cigarettes did you usually smoke in
             one day? ___ cigarettes per day
   □ Yes → c. Are you still smoking?
     □ No → How old were you when you stopped smoking? ___ years old
     □ Yes

17. “Second-hand” smoke exposure occurs when you are in close enough to a smoker to smell the
tobacco smoke. During the past month, how many hours per week were you exposed to second
hand smoke?
   □ 0 or < 1 hour
   □ 1-4 hours
   □ 5-9 hour
   □ 10-14 hours
   □ 15 or more hours
18. a. How often did you drink alcohol beverages during the past month?
   
   *1 alcohol beverage = 12oz beer (1 bottle), 5oz wine (1 glass) or 1.5oz liquor*
   
   □ More than once a day
   □ About every day
   □ 4 to 5 times a week
   □ 2 to 3 times a week
   □ Once a week
   □ 2 to 3 times a month
   □ Once a month
   □ Less than once a month

   b. On those days when you drank alcohol, how many drinks did you have?
      ___ drink(s)

PHYSICAL ACTIVITY

We are interested in finding out about the kinds of physical activity that you have done in a *usual week* over the *past month.*

Think about all the *vigorous* activities that you did in a usual week in the past month. *Vigorous* physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. Think *only* about those physical activities that you did for at least 10 minutes at a time.

19. In a usual week in the past month, on how many days did you do vigorous physical activities like heavy lifting, digging, aerobics, or fast bicycling?
   
   ___ days per week
   
   □ No vigorous physical activities ➔ Skip to question 21

20. How much time did you usually spend doing vigorous physical activities on one of those days?
   
   ___ hours per day
   ___ minutes per day
Think about all the **moderate** activities that you did in an average week in the past month. **Moderate** activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

21. During a usual week in the past month, on how many days did you do **moderate** physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.
   
   _______ days per week

   [ ] No moderate physical activities   ➔ Skip to question 23

22. How much time did you usually spend doing moderate physical activities on one of those days?

   _______ hours per day

   _______ minutes per day

Think about the time you spent walking in a usual week in the past month. This includes at work and at home, walking to travel from place to place, and any other walking that you might do solely for recreation, sport, exercise, or leisure.

23. During a usual week in the past month, on how many days did you walk for at least 10 minutes at a time?

   _______ days per week

   [ ] No walking   ➔ Skip to question 25

133
24. How much time did you usually spend walking on one of those days?
   _____ hours per day
   _____ minutes per day

The last question is about the time you spent sitting on weekdays during a usual week in the past month. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

25. During a usual week in the past month, how much time did you spend sitting on a week day?
   _____ hours per day
   _____ minutes per day
Appendix D

Model Diagnostics

In addition to checking for unusual and influential points residual plots were generated to test the linear regression assumptions of homoscedasticity and normality. From the plot of residual versus predicted values (Figure D-1), a pattern of greater dispersion in data points towards the right end of the graph can be seen. This is suggestive of mild heteroscedasticity. However, the p-value ($\chi^2=77.8$, $p=0.92$) calculated from the White test for heteroscedasticity indicates that the null hypothesis of homogeneity cannot be rejected. The kernel density plot (Figure D-2), and residual density and probability plots (Figure D-3) demonstrate that the data follows a nonlinear pattern and suggest a residual distribution that is right skewed.
Figure D-1. Test for heteroscedasticity showing a plot of residual versus predicted values.
Figure D-2. Test for normality showing a kernel density plot of the model residual distribution.
Figure D-3. Residual density and probability plots. A histogram of studentized residuals is shown in the top panel and a Q-Q plot is shown by the bottom panel.
Sensitivity Analysis

As the residuals were not normally distributed in this study population a sensitivity analysis with natural log-transformed rLTL was performed (Table D-1). The results demonstrate that the multiple linear regression models with untransformed and natural log-transformed rLTL are qualitatively the same.
Table D-1. Results from multivariable linear regression with natural log-transformed rLTL presenting fully adjusted regression coefficients relating lifestyle exposure variables to rLTL.

<table>
<thead>
<tr>
<th>Lifestyle exposure</th>
<th>Log-transformed rLTL a</th>
<th>N (%)</th>
<th>Coefficient ± SE</th>
<th>p-value b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol consumption</td>
<td></td>
<td></td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>Abstainer</td>
<td></td>
<td>45 (9)</td>
<td>-0.014 ± 0.050</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>325 (68)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td>68 (14)</td>
<td>-0.078 ± 0.043</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>39 (8)</td>
<td>0.044 ± 0.055</td>
<td></td>
</tr>
<tr>
<td>p-trend</td>
<td></td>
<td></td>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>Never</td>
<td></td>
<td>325 (68)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td></td>
<td>90 (19)</td>
<td>-0.076 ± 0.039</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td></td>
<td>62 (13)</td>
<td>-0.091 ± 0.045*</td>
<td></td>
</tr>
<tr>
<td>Pack-years smoking</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Never smoker</td>
<td></td>
<td>326 (68)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Lower tertile (&gt;0 – 6.25)</td>
<td></td>
<td>66 (14)</td>
<td>-0.10 ± 0.043*</td>
<td></td>
</tr>
<tr>
<td>Middle tertile (&gt;6.25 – 16)</td>
<td></td>
<td>43 (9)</td>
<td>-0.14 ± 0.053**</td>
<td></td>
</tr>
<tr>
<td>Higher tertile (&gt;16)</td>
<td></td>
<td>42 (9)</td>
<td>0.016 ± 0.055</td>
<td></td>
</tr>
<tr>
<td>p-trend</td>
<td></td>
<td></td>
<td></td>
<td>0.14</td>
</tr>
<tr>
<td>Total physical activity</td>
<td></td>
<td></td>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>42 (9)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td>225 (47)</td>
<td>0.037 ± 0.054</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>210 (44)</td>
<td>0.080 ± 0.055</td>
<td></td>
</tr>
<tr>
<td>p-trend</td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>Vigorous physical activity (MET-mins)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Q1 (0)</td>
<td></td>
<td>147 (31)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>Q2 (&gt;0-&lt;840)</td>
<td></td>
<td>91 (19)</td>
<td>0.038 ± 0.042</td>
<td></td>
</tr>
<tr>
<td>Q3 (840-&lt;1920)</td>
<td></td>
<td>134 (28)</td>
<td>-0.0005 ± 0.038</td>
<td></td>
</tr>
<tr>
<td>Q4 (≥1920)</td>
<td></td>
<td>103 (22)</td>
<td>0.13 ± 0.042**</td>
<td></td>
</tr>
<tr>
<td>p-trend</td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
</tbody>
</table>

Abbreviations: SE=Standard error; Ref.=Referent
Adjusted for age, sex, and lipid ratio while also adjusting for the other lifestyle exposures of interest. For example, the model with alcohol consumption adjusts for age, sex, lipid ratio, smoking status, and total physical activity.

Overall p-value computed by F-statistic.

p<0.05 in comparison to referent computed by t-statistic

** p<0.01 in comparison to referent computed by t-statistic
Appendix E

Detectable Difference Calculations

Table E-1. Detectable difference calculation for the association between telomere length and alcohol consumption, smoking and physical activity. Telomere length was standardized to a standard normal distribution with a mean of 0 and standard deviation of 1. Detectable difference calculations were performed using PS: Power and Sample Size Calculation version 3.0.34, 2011 for Mac (1).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha (α)</td>
<td>0.05</td>
</tr>
<tr>
<td>Sample size (N)</td>
<td>600</td>
</tr>
<tr>
<td>Power (1-β)</td>
<td>0.80</td>
</tr>
</tbody>
</table>

**Alcohol consumption**

<table>
<thead>
<tr>
<th>Prevalence of exposure (%)</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>n unexposed (Low)</td>
<td>463</td>
</tr>
<tr>
<td>n exposed (High)</td>
<td>56</td>
</tr>
<tr>
<td>r</td>
<td>7.20</td>
</tr>
<tr>
<td>Minimum detectable difference</td>
<td><strong>0.40 of a standard deviation</strong></td>
</tr>
</tbody>
</table>

**Cigarette smoking**

<table>
<thead>
<tr>
<th>Prevalence of exposure (%)</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>n unexposed (Never)</td>
<td>404</td>
</tr>
<tr>
<td>n exposed (Current)</td>
<td>80</td>
</tr>
<tr>
<td>r</td>
<td>5.05</td>
</tr>
<tr>
<td>Minimum detectable difference</td>
<td><strong>0.34 of a standard deviation</strong></td>
</tr>
</tbody>
</table>

**Physical activity**

<table>
<thead>
<tr>
<th>Prevalence of exposure (%)</th>
<th>44</th>
</tr>
</thead>
<tbody>
<tr>
<td>n unexposed (Low)</td>
<td>54</td>
</tr>
<tr>
<td>n exposed (High)</td>
<td>264</td>
</tr>
<tr>
<td>r</td>
<td>0.20</td>
</tr>
<tr>
<td>Minimum detectable difference</td>
<td><strong>0.42 of a standard deviation</strong></td>
</tr>
</tbody>
</table>
**Table E-2.** Detectable effect calculation within the CYP2E1 wild-type and variant genotype strata.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha (α)</td>
<td>0.05</td>
</tr>
<tr>
<td>Sample size (N)</td>
<td>600</td>
</tr>
<tr>
<td>Power (1-β)</td>
<td>0.20</td>
</tr>
<tr>
<td>Standard Deviation (σ)</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>0</td>
</tr>
</tbody>
</table>

**Wild-type (*1A/*1A)**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>540</td>
</tr>
<tr>
<td>Prevalence of exposure (%)</td>
<td>22</td>
</tr>
<tr>
<td>n unexposed (≤ 7 drinks/week)</td>
<td>421</td>
</tr>
<tr>
<td>n exposed (&gt;7 drinks/week)</td>
<td>119</td>
</tr>
<tr>
<td>r</td>
<td>3.54</td>
</tr>
<tr>
<td>Detectable difference</td>
<td><strong>0.29 of a standard deviation</strong></td>
</tr>
</tbody>
</table>

**Variant (*1A/*5B and *5B/*5B)**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>61</td>
</tr>
<tr>
<td>Prevalence of exposure (%)</td>
<td>22</td>
</tr>
<tr>
<td>n unexposed (≤ 7 drinks/week)</td>
<td>48</td>
</tr>
<tr>
<td>n exposed (&gt;7 drinks/week)</td>
<td>13</td>
</tr>
<tr>
<td>r</td>
<td>3.69</td>
</tr>
<tr>
<td>Detectable difference</td>
<td><strong>0.89 of a standard deviation</strong></td>
</tr>
</tbody>
</table>