EFFECT OF ALCOHOL CONSUMPTION AND ALCOHOL DEHYDROGENASE 1C (ADH1C) POLYMORPHISMS ON TOTAL PLASMA HOMOCYSTEINE LEVELS

By

Dana Al-Bargash

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

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

Copyright © Dana Al-Bargash, 2008
Abstract

BACKGROUND: Evidence supports that an elevated level of total plasma homocysteine (tHcy) signifies a breakdown in the methionine-homocysteine cycle. This may result in folate deficiency, DNA methylation and oxidative stress, all of which are potential mechanisms that may lead to cancer. Few studies examined the effects of alcohol consumption on tHcy levels. No studies considered polymorphisms in alcohol dehydrogenase 1C (ADH1C), a gene that encodes for an enzyme that metabolizes alcohol to acetaldehyde, a folate antagonist; ADH1C*1 encodes for an enzyme with a higher capacity to generate acetaldehyde than ADH1C*2.

PURPOSE: This study examined the association between alcohol intake and risk of elevated tHcy while exploring a potential gene-environment interaction with polymorphisms in ADH1C.

METHOD: This was a case-control study nested in a larger cross-sectional study funded by Canadian Institutes of Health Research. The target population was recruited from Kingston and Halifax from 2006 to 2008 and included 100 cases and 187 controls selected from healthy male and female subjects aged 20-50 years. Cases were defined as subjects with tHcy ≥ 10 μmol/L and controls < 10 μmol/L. Alcohol consumption was categorized into three groups: ≤12.0 g/day, 12.1–24.0 g/day, and >24.0 g/day. ADH1C was dichotomized by collapsing ADH1C*1/*2 and ADH1C*2/*2 into one group and examined against ADH1C*1/*1.
RESULTS: Compared to ≤12.0 g/day, odds ratio (OR) for 12.1-24.0 g/day was 0.53 [95% confidence interval (CI), 0.25-1.13] and for >24.0 g/day was 1.19 (95% CI, 0.60-2.24), suggesting a J-shaped trend with risk of elevated tHcy. A reduced OR was observed for ADH1C*1/*1 (OR= 0.52, 95% CI, 0.27-1.03). The alcohol-ADH1C interaction was not statistically significant (p-value = 0.21), though a stronger J-shape trend was suggested in ADH1C*1/*1. Among consumers of ≤12.0 g/day, a reduced measure of effect was observed for ADH1C*1/*1 (OR= 0.44, 95% CI, 0.19-1.00).

CONCLUSION: A J-shaped trend was suspected between risk of elevated tHcy and alcohol consumption. Additionally, a nonsignificant reduced effect of ADH1C*1/*1 on risk of elevated tHcy, with a more pronounced effect in the lowest group of alcohol consumption. This suggests that ADH1C may be associated to homocysteine through factors unrelated to alcohol intake.
Co-Authorship Statement

Dana Al-Bargash, the student investigator, was involved in composing the research questions for this thesis, identifying the study design, locating laboratory protocols for genotyping data, undertaking all PCR-RFLP required for genotyping, managing collected data, analyzing results, interpreting results, writing, and editing the manuscript.

The original larger study from which data used in this thesis was collected was designed and implemented in collaboration between Dr. Will King from the department of Community Health and Epidemiology at Queen’s University who served as the Principal Investigator, Dr. Thomas Massey from the department of Toxicology and Pharmacology at Queen’s University, Dr. Dodds from Dalhousie, and S. Perkins from the University of Ottawa. Dr. Will King oversaw the design of this research study, its implementations, interpretations and representation of results. He also assisted in planning and editing the manuscript. Dr. Thomas Massey was involved in planning of this study, while providing assistance with laboratory methods and interpretations of ideas in molecular biology. He also assisted in editing the manuscript.

Canadian Institute of Health Research (CIHR) was the primary source of financial assistance required for this research.
Acknowledgements

There are many people I would like to send my deepest gratitude but I will keep my list as short as possible. I would first like to thank my primary supervisor, Dr. Will King for his invaluable source of support and guidance from the beginning of my study to the final stages. A student could not ask for a more patient and committed supervisor who was always there to make some sense of much confusion I often found myself confronted with. I would also like to express my sincerest gratitude to Dr. Thomas Massey for providing me with feedback and insights during my many stages of troubleshooting in the lab and for his constructive comments and valuable suggestions that helped me make the most of my transdisciplinary experience.

I’d like to acknowledge the staff and faculty of the Department of Community Health and Epidemiology, for without their help and guidance, I could not have come this far. I would like to acknowledge CIHR Transdisciplinary Training Program in Cancer Research for providing me with a transdisciplinary

I also would like to cordially thank Vikki, Gwyneth, Janet, Em, Sarah, and Marianne and all the great friends I made in the last two years for their kind support and assistance, and whose presence made my journey over the last two years a pleasant one

Finally I would like to sincerely thank my parents for their endless support and love and for teaching me that I can accomplish anything I set out to do; I would like to also acknowledge my brother and sister who always gave me their undivided attention.
Table of Content

Abstract.................................................................................................................................i
Co-Authorship Statement .................................................................................................iii
Acknowledgements ............................................................................................................iv
Table of Content.................................................................................................................. v
List of Tables ......................................................................................................................ix
List of Figures ................................................................................................................... x
List of Abbreviation...........................................................................................................xi

CHAPTER 1: INTRODUCTION...................................................................................... 1

1.1 General introduction ............................................................................................... 1
1.2 Objectives ............................................................................................................... 3
1.3 Context of research ................................................................................................. 3
1.4 Thesis outline ........................................................................................................... 4
1.5 References............................................................................................................... 6

CHAPTER 2: LITERATURE REVIEW.......................................................................... 7

2.1 Alcohol consumption ............................................................................................. 7
2.1.1 Definition and properties ............................................................................. 7
2.1.2 Alcohol absorption and elimination ............................................................. 7
2.1.2.1 Pathways of alcohol metabolism ...................................................... 8
2.1.2.1.1 ADH – ALDH pathway ................................................. 8
2.1.2.1.2 CYP2E1 pathway ................................................................ 11
2.1.2.1.3 Catalase pathway ................................................................ 11
2.1.3 Unpleasant effects of alcohol consumption ............................................... 12
2.1.3.1 Literature review between alcohol consumption and cancer .......... 12
2.1.3.2 Possible mechanisms of alcohol-mediated carcinogenesis ............. 13
2.2 Polymorphic enzymes affecting the rate of alcohol metabolism ...................... 14
2.2.1 ADH1C enzyme ......................................................................................... 14
2.2.1.1 ADH1C polymorphisms ................................................................ 14
2.2.1.2 ADH1C genotype frequencies ....................................................... 15
2.2.1.3 Roles of ADH1C ....................................................................... 16
2.2.1.4 ADH1C and cancer ..................................................................... 16
2.2.1.4.1 ADH1C and head and neck cancer ....................................... 16
2.2.2 ADH1B enzyme ......................................................................................... 19
2.2.2.1 ADH1B polymorphisms ............................................................. 19
2.2.2.2 ADH1B genotype frequencies ....................................................... 19
2.2.3 ALDH2 enzyme ......................................................................................... 19
2.2.3.1 ALDH2 polymorphisms ............................................................. 19
2.2.3.2 ALDH2 genotype frequencies ....................................................... 20
2.2.4 CYP2E1 enzyme ......................................................................................... 20
2.2.4.1 CYP2E1 polymorphisms ............................................................. 20
2.2.4.2 CYP2E1 genotype frequencies ....................................................... 21
2.3 Homocysteine ....................................................................................................... 22
CHAPTER 2: HOMOCYSTEINE AND ADVERSE HEALTH OUTCOMES ............................................. 22
2.3.1 Definition and properties of homocysteine ................................................................. 22
2.3.2 Determinants of homocysteine ..................................................................................... 24
2.3.3 Homocysteine and adverse health outcomes .............................................................. 25
2.3.4 Biomarkers .................................................................................................................... 26
2.3.5 Desirable homocysteine levels .................................................................................... 27
2.3.6 Stability of tHcy over time ........................................................................................... 28
2.3.7 Biological mechanism of homocysteine-mediated carcinogenesis ......................... 29
2.4 Alcohol consumption and homocysteine ..................................................................... 31
2.4.1 Literature ....................................................................................................................... 31
2.4.2 Biological effects of alcohol on methionine-homocysteine cycle ............................. 33
2.5 ADH1C polymorphisms and tHcy ................................................................................ 36
2.6 References ....................................................................................................................... 37

CHAPTER 3: STUDY DESIGN AND METHODS ................................................................. 48
3.1 Study design .................................................................................................................... 48
3.2 Source population and subjects ..................................................................................... 49
3.2.1 Cases and controls ....................................................................................................... 49
3.2.2 Ethnicity ....................................................................................................................... 50
3.3 Data collection ............................................................................................................... 51
3.3.1 Subject recruitment ..................................................................................................... 51
3.3.2 Blood extraction and ADH1C genotyping ................................................................. 52
3.3.3 Measurement of alcohol consumption ...................................................................... 54
3.3.4 Other variables ............................................................................................................ 55
3.4 Data validity .................................................................................................................... 56
3.4.1 ADH1C genotypes ...................................................................................................... 56
3.4.2 Measurement of alcohol consumption ................................................................... 57
3.5 Data management .......................................................................................................... 58
3.6 Sample size .................................................................................................................... 58
3.7 Detectable effect ............................................................................................................. 59
3.8 Design effect ................................................................................................................. 62
3.9 Data analysis ................................................................................................................ 62
3.9.1 Preliminary analysis .................................................................................................. 62
3.9.2 Model selection and control for confounding ........................................................... 64
3.9.3 Main analyses .......................................................................................................... 64
3.9.3.1 Effects of alcohol consumption ......................................................................... 64
3.9.3.2 Effects of genotype ......................................................................................... 65
3.9.3.3 Gene-environment interaction ........................................................................ 65
3.10 Ethical considerations ................................................................................................. 66
3.11 References ................................................................................................................... 67

CHAPTER 4: MANUSCRIPT ......................................................................................... 70
SUMMARY ......................................................................................................................... 71
INTRODUCTION ................................................................................................................ 73
METHODS ......................................................................................................................... 75
Study population .............................................................................................................. 75
CHAPTER 5: SUPPLEMENTARY RESULTS ......................................................... 89
  5.1 Introduction........................................................................................................... 89
  5.2 Distributions of variables...................................................................................... 89
    5.2.1 Distribution of alcohol ............................................................................... 89
    5.2.2 Distribution of ADH1C Genotypes .............................................................. 91
    5.2.3 Distribution of tHcy ................................................................................... 92
  5.3 Assessment of potential for confounding ............................................................. 93
    5.3.1 Relationships between alcohol, ADH1C and potential confounders .......... 94
    5.3.2 Relationship between tHcy and potential confounders.............................. 95
    5.3.3 Selection of variables into the model ......................................................... 98
  5.4 Model fit and goodness of fit ............................................................................. 98
  5.5 Main analysis ........................................................................................................ 99
    5.5.2 Serum folate ............................................................................................... 99
    5.5.3 Sex............................................................................................................ 101
  5.6 Sensitivity analyses............................................................................................. 102
    5.6.1 Alcohol intake: 3 consumption groups versus 4 consumption groups .... 102
    5.6.2 ADH1C: 2 genotype groups versus 3 genotype groups ......................... 103
    5.6.3 Investigating tHcy as continuous outcome .............................................. 103
    5.6.4 Removing serum folate from the model .................................................. 104
    5.6.5 Results of design effect ........................................................................... 105
  5.7 Validity of genotyping techniques ...................................................................... 106
  5.8 Validity of genotyping results............................................................................. 106
  5.9 Odds ratio (OR) as an estimate of relative risk (RR) .......................................... 107
  5.10 References......................................................................................................... 109

CHAPTER 6: GENERAL DISCUSSION .................................................................... 110
  6.1 Summary of key findings.................................................................................... 110
    6.1.1 Main results .............................................................................................. 110
  6.2 Methodological considerations ........................................................................... 110
    6.2.1 Outcome assessment ............................................................................... 110
    6.2.2 Genotype assessment ............................................................................. 111
    6.2.3 Exposure assessment .............................................................................. 111
    6.2.4 Study design ........................................................................................... 114
    6.2.5 Confounders .......................................................................................... 115
      6.2.5.1 Serum folate .................................................................................. 115
    6.2.6 Sample size .............................................................................................. 116
  6.3 Alcohol and carcinogenesis of the colon ............................................................ 116
  6.4 Generalizability ................................................................................................... 119
  6.5 Conclusion .......................................................................................................... 120
  6.6 Implications......................................................................................................... 120
6.7 Future Direction ........................................................................................................ 121
6.7 References............................................................................................................. 122

APPENDIX A .................................................................................................................. 125
APPENDIX B .................................................................................................................. 126
List of Tables

Chapter 3

Table 3.1: Detectable effect calculations for alcohol consumption.................................60
Table 3.2: Detectable effect calculations for ADH1C genotypes......................................61
Table 3.3: Detectable effect calculations for ADH1C genotypes......................................61

Chapter 4

Table 4.1: Characteristics of study population...............................................................80
Table 4.2: Alcohol consumption, ADH1C polymorphisms and cases of elevated tHcy........81
Table 4.3: Interaction between alcohol consumption and ADH1C polymorphisms...........81

Chapter 5

Table 5.1: Relationship between covariates.................................................................96
Table 5.2: Illustration of variables in continuous and categorical forms............................97
Table 5.3: Interaction between serum folate and alcohol consumption..........................100
Table 5.4: Interaction between serum folate status and ADH1C polymorphisms...............101
Table 5.5: Interaction between sex and ADH1C polymorphisms..................................102
Table 5.6: Sensitivity analysis on categorization of alcohol consumption.......................102
Table 5.7: Sensitivity analysis on categorization of ADH1C genotypes............................103
Table 5.8: Linear regression analysis............................................................................104
Table 5.9: Sensitivity analysis after removing serum folate from model..........................105
Table 5.10: Sensitivity analysis including random effect................................................106
Table 5.11: Estimating relative risk..............................................................................108
List of Figures

Chapter 2

Figure 2.1: Chemical structure of alcohol ..................................................................................7
Figure 2.2: Alcohol metabolism: ADH-ALDH Pathway .........................................................10
Figure 2.3: Chemical structure of homocysteine .................................................................22
Figure 2.4: Methionine-homocysteine cycle ........................................................................24
Figure 2.5: J-Shape Trend Between Alcohol consumption and tHcy levels .......................32
Figure 2.6: Effects of Acetaldehyde on Methionine-Homocysteine Cycle .........................35

Chapter 4

Figure 4.1: ADH1C genotyping by PCR-RFLP ...................................................................78

Chapter 5

Figure 5.1: Distribution of alcohol consumption .................................................................91
Figure 5.2: Distribution of homocysteine ..........................................................................93
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADH1C</td>
<td>Alcohol Dehydrogenase Type 1C</td>
</tr>
<tr>
<td>ADH1C*1</td>
<td>Fast alcohol metabolizer variant allele; translates Isoleucine amino acid</td>
</tr>
<tr>
<td>ADH1C*2</td>
<td>Slow alcohol metabolizer variant allele; translates Valine amino acid</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 Pairs</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CYP2E1</td>
<td>Cytochrome P450 2E1</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>Ile349Val</td>
<td>Isoleucine to Valine Amino Acid Substitution at codon 349</td>
</tr>
<tr>
<td>IPAQ</td>
<td>International Physical Activity Questionnaire</td>
</tr>
<tr>
<td>MAT</td>
<td>Methionine Adenosyltransferase reaction</td>
</tr>
<tr>
<td>MEOS</td>
<td>Microsomal Ethanol Oxidizing System</td>
</tr>
<tr>
<td>MET</td>
<td>Metabolic Equivalent</td>
</tr>
<tr>
<td>MS</td>
<td>Methionine Synthase</td>
</tr>
<tr>
<td>MTHFR</td>
<td>Methylene tetrahydrofolate reductase</td>
</tr>
<tr>
<td>NADH</td>
<td>Reduced Nicotinamide Adenine Dinucleotide</td>
</tr>
<tr>
<td>NAD+</td>
<td>Nicotinamide Adenine Dinucleotide</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>QF</td>
<td>Quantity-Frequency</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction Fragment Length Polymorphism</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive Oxygen Species</td>
</tr>
<tr>
<td>SAH</td>
<td>S-adenosyl-homocysteine</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosyl-methionine</td>
</tr>
<tr>
<td>Std Dev</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>tHcy</td>
<td>Total Plasma Homocysteine</td>
</tr>
<tr>
<td>THF</td>
<td>Tetrahydrofolate</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

1.1 General introduction

Cancer is a significant public health problem and is the leading cause of premature deaths in Canada. An estimated 166,400 incident cases and 73,800 deaths from cancer will occur in Canada alone in 2008 (1). There is sufficient evidence that modifiable, environmental, lifestyle, and genetic factors have the ability to impact the risk of cancer. Despite the copious attention this multi-factorial disease has received over the years, its etiology remains unclear. It is therefore important to identify modifiable risk factors as well as expand our knowledge of the mechanisms underlying carcinogenesis in order for prevention strategies can be implemented.

“Alcohol” is a term that is used interchangeably with “ethanol” in health research; this thesis has chosen to consistently refer to “alcohol”. A growing body of evidence has supported a link between heavy alcohol consumption and high risk of cancer at several sites. An estimated of 2 to 4 % of all cases of cancer are thought to be caused by alcohol consumption (2). Chronic alcohol consumption has been established as a risk factor for cancer of the liver, the upper respiratory, and digestive tracts including mouth, pharynx, larynx, and esophagus (3); it is suspected as a probable agent for increase in risk of breast and colorectal cancer (4,5,6).

Multiple mechanisms are involved in alcohol-associated development of cancer. Recently, research has focused on the mutagenic and carcinogenic downstream metabolite of alcohol oxidation, acetaldehyde (3). Alcohol and acetaldehyde are suspected to disrupt the methionine-homocysteine cycle, a cycle that is implicated to
influence risk of colorectal cancer (7). This disruption, particularly through deficiency of either vitamin B₁₂ (cobalamin) or folate (used interchangeably with folic acid), has been shown to promote carcinogenesis by interference with DNA (deoxyribonucleic acid) synthesis, repair, and methylation (7). The status of this methionine-homocysteine cycle is often reflected by concentrations of total plasma homocysteine (tHcy), where elevated levels in the blood signal a disruption in this vital cycle, resulting in far-reaching biochemical and life consequences (8). In fact, a disruption in the methionine-homocysteine cycle is postulated to promote the production of reactive oxygen species (ROS) and altered methylation capacity, which have shown oncogenic potential (9).

In humans, variation in genes that encode for several alcohol metabolizing enzymes can affect the rate of alcohol metabolism. It is estimated that over 80% of alcohol is oxidized to acetaldehyde by *Alcohol Dehydrogenase (ADH)* (10). Variants in enzymes that affect the conversion rate of alcohol to acetaldehyde have been described for two classes of ADH enzymes, ADH2 and ADH3, where their nomenclature has recently been changed to ADH1B and ADH1C, respectively (10). The goal for this research is to examine the association between alcohol consumption and tHcy while considering polymorphisms in *ADH1C* for potential modification of this relationship.
1.2 Objectives

This thesis examines the relationship between alcohol consumption and risk of elevated tHcy (where subjects with ≥ 10 µmol/L are identified as cases). A J-shaped trend is hypothesized to exist between the volume of alcohol consumption and risk of elevated tHcy; nondrinkers, light, and heavy consumers of alcohol are suspected to be at higher risk of elevated tHcy compared to moderate consumers of alcohol.

The alcohol-tHcy relationship is postulated to be modified by an individual’s ability to metabolize alcohol. To address this, polymorphisms in the candidate gene encoding for ADH1C enzyme are investigated in combination with alcohol consumption to identify if genetic susceptibility exists. A polymorphism is an identifiable variant form of a gene that occurs in at least 1% of the population (10). It is hypothesized that the relationship between alcohol consumption and risk of elevated tHcy level is modified by ADH1C genotypes. In particular, the disruptive relationship between alcohol consumption and tHcy is suspected to be stronger in ADH1C*1/*1.

1.3 Context of research

The ideas of this study including the assessment of a gene-environment interaction between alcohol consumption and ADH1C polymorphisms were formulated by Dana Al-Bargash, the student researcher. The student researcher uncovered objectives and devised a plan of research and established the study design. All required laboratory and epidemiological analyses and interpretations were conducted by the student researcher. This study was conducted in a cancer-related setting, and was nested within a larger CIHR funded cross-sectional study. The larger study provided information on alcohol
CHAPTER 1: INTRODUCTION

consumption, additional variables, and the outcome of interest; the student researcher established the required indices to undertake the research questions. Student funding and transdisciplinary training was provided by CIHR transdisciplinary training program in cancer research.

1.4 Thesis outline

The second chapter of this thesis presents a literature review of epidemiological evidence of the methionine-homocysteine cycle and potential roles in carcinogenesis. The chapter also delineates suspected effects of alcohol consumption on this cycle with supporting evidence for its participation in mechanisms underlying carcinogenesis. Additionally, the potential role of genetic susceptibility between alcohol consumption, \textit{ADH1C} polymorphisms, and risk of elevated tHcy will be outlined followed by plausible biological interpretations.

The third chapter of this thesis describes the stages taken to conduct this study. This includes methodological issues as well as study methods that are not included in the manuscript. This chapter also provides a summary of the study population, methods of confounder selection, methods used to assess data validity, sensitivity analyses, measures of detectable effects, and statistical analyses of additional data.

The fourth chapter of this thesis is a manuscript that is formatted for submission as a research article to \textit{International Journal of Epidemiology}. The manuscript describes the case-control study design nested within a larger cross-sectional study. This smaller study
investigates the effects of alcohol consumption, *ADH1C* (alcohol dehydrogenase class IC) polymorphisms and associations to the risk of elevated tHcy.

The fifth chapter of this thesis presents additional results that were not discussed in chapter 4, including additional analyses, study validity and sensitivity tests.

The sixth and final chapter of this thesis presents a general discussion of the study objectives, findings, strengths, limitations, conclusion and future implications.


2.1 Alcohol consumption

2.1.1 Definition and properties

Alcohol (CH₃CH₂OH) is a popular mood-altering beverage that is likely to be the most widely investigated drug in the world due to its widespread abuse and unique pharmacological properties (Figure 2.1) (1). This non-essential substance is a hydrophilic and normally found in liquid form at room temperature (25°C).

![Chemical structure of alcohol](image)

**Figure 2.1** Chemical structure of alcohol

2.1.2 Alcohol absorption and elimination

After oral administration, alcohol is readily absorbed into the bloodstream via the gastrointestinal tract starting from the mouth to the rectum and rapidly distributed among various tissues of the body leading to a uniform distribution in the total body water (2). It begins with absorption of about 20% of alcohol through the stomach wall and the remaining 80% through the duodenum and small intestine (2). Small amounts of alcohol are initially metabolized in the stomach. Alcohol then passes into the portal vein (a large vein that carries blood from the digestive tract to the liver) into the liver where it soon reaches its maximum rate of metabolism, which is defined as the hepatic first-pass metabolism (1,2). Alcohol is primarily metabolized in the liver (92–95%), but is also
removed and metabolized in extrahepatic tissues such as the stomach, small intestine, colon, kidneys, and lungs (2,3). Overall, 95-98% of ingested alcohol is metabolized. Alcohol can be eliminated by urine (0.3%) either unchanged or through conjugation with glucuronic acid and ethyl glucuronide (2). Additional pathways of elimination of unchanged alcohol occur through excretion in the breath (0.7%), and sweat (0.1%) (2).

The liver can process approximately ½ ounce (1 drink) of alcohol every hour; however, a considerable degree of interindividual and ethnic variability in rate of metabolism exists. The amount of metabolizing enzymes present in the liver, which vary between individuals, and polymorphisms in particular genes involved in the elimination of alcohol predict the rate at which alcohol is metabolized (4,5). The rate of alcohol metabolism also depends on other modifiable and non-modifiable factors such as chronic alcohol consumption, recent food intake, gender, body weight, age, smoking, time of day, and rate of consumption (5,6).

2.1.2.1 Pathways of alcohol metabolism

Several pathways metabolize alcohol.

2.1.2.1.1 ADH – ALDH pathway

The most common pathway of alcohol metabolism is via a two-step pathway through acetaldehyde to acetate (Figure 2.2) (3). Most (80%) of ingested alcohol is first metabolized to acetaldehyde in a reversible reaction mainly by an enzyme present in the cytosol fraction of cells, alcohol dehydrogenase (ADH), followed by a non-reversible metabolism of acetaldehyde to acetate by aldehyde dehydrogenase (ALDH) (1-3).
The ADH protein is a dimer of 40,000-Da subunits that is divided into five classes based on amino acid sequence, which are building blocks of proteins, and structural similarities (1). The human liver contains several classes of ADH enzymes (1,2). Class I corresponds to the typical liver enzymes that are derived from three separate genes located on chromosome 4. These genes produce α, β, and γ subunits, which combine in an assortment of dimeric combinations to form class I ADH isoenzymes (1). The alleles located at the \textit{ADH1A}, \textit{ADH1B}, and \textit{ADH1C} loci correspond to α, β, and γ subunits. These three class I enzymes share 93% sequence identity but differ in their substrate specificity and development expression (5,6). Of all classes of ADH, only \textit{ADH1B} and \textit{ADH1C} exhibit genetic variation in particular ethnic populations (7-9).

In humans, there are at least four classes of the ALDH isoenzymes. It is aldehyde dehydrogenase 2 (ALDH2), an enzyme that is encoded by a polymorphic gene, which eliminates majority of the acetaldehyde generated during alcohol metabolism (5,10).

The ADH reaction is readily reversible and is subject to product inhibition by acetaldehyde. Therefore, changes in the activity of ADH can cause changes in the rate of alcohol metabolism. However, acetaldehyde does not accumulate to high concentrations in the liver as it is rapidly and irreversibly oxidized to acetate by aldehyde dehydrogenase (ALDH) (3). This is not always the case as some populations lack ALDH activity (discussed in section 2.2.3) (3). Acetate enters the normal metabolic pools, mainly extrahepatic, where it is ultimately converted to carbon dioxide (CO$_2$) and water (H$_2$O) (1).
The conversion of alcohol to acetaldehyde and acetaldehyde to acetate results in the generation of NADH (reduced nicotinamide adenine dinucleotide) from NAD+ (nicotinamide adenine dinucleotide) (Figure 2.2). NADH, in turn, participates in many essential biochemical reactions in the cell (11). Generally, there is a limit to the rate at which NADH can be converted back to NAD+, which is required for the continuous oxidation of alcohol and acetaldehyde molecules. In fact, a high rate of alcohol metabolism generates excess amounts of NADH, causing the cell to no longer maintain the normal NAD+/NADH ratio. Thus accumulation in acetaldehyde, which exerts strong product inhibition on ADH reaction, and the excess in NADH will poise the reaction toward formation of alcohol and NAD+ (12).

![Figure 2.2 Alcohol metabolism: ADH-ALDH Pathway](image)

The rate of ADH activity can vary depending on the concentrations of alcohol, acetaldehyde, NAD+, and NADH, where the latter three are also regulated by the rate of activity of ADH and ALDH. The first pass-metabolism, which could predict systemic availability of alcohol, is affected by the gastric and hepatic ADH (2). Concentrations of acetaldehyde are regulated by the balance between the activities of ADH and ALDH in
CHAPTER 2: LITERATURE REVIEW

The rate-limiting step of alcohol metabolism is the oxidation of alcohol by ADH (7). ADH has an important function in limiting the accumulation of acetaldehyde since the equilibrium constant for alcohol to acetaldehyde is in the direction of alcohol (12).

2.1.2.1.2 CYP2E1 pathway

Another pathway through which alcohol is metabolized is the Microsomal Ethanol Oxidizing Systems (MEOS), which utilizes the cytochrome P450 2E1 (CYP2E1) enzyme (1). CYP2E1 is predominantly induced in the hepatocyte by high concentrations of alcohol; it is also expressed in the heart, brain, lung, and neutrophils and macrophages of the immune system. This enzyme is a source of reactive oxygen species, leading to liver injury (13). Like ADH, CYP2E1 oxidizes alcohol and generates acetaldehyde. In moderate consumption, the CYP2E1 enzyme is responsible for approximately 2-3% of alcohol metabolism, having negligible effects on rate of clearance of alcohol in the blood. However, it can be induced by chronic administration of alcohol where its activity can increase 20-fold, accounting for 10-20% of alcohol metabolism (13). Authors suggested that CYP2E1 normally accounts for 3-7% of alcohol removal, having little effect on rate of alcohol clearance at moderate intake.

2.1.2.1.3 Catalase pathway

Catalase is another enzyme that metabolizes alcohol to acetaldehyde. This protein is expressed in all tissues with highest levels present in the liver, kidney, heart, and red blood cells (10,14). In the presence of an H₂O₂ generating system, catalase is able to oxidize alcohol in vitro generating acetaldehyde and water. The contribution of catalase
to the systemic elimination of alcohol *in vivo* remains controversial, however a consensus was reached in that this enzyme plays no major role in alcohol metabolism (10). Metabolism by catalase *in vivo* is unlikely as it lacks hydrogen peroxide that is necessary for this reaction to proceed (12).

### 2.1.3 Unpleasant effects of alcohol consumption

Some side effects related to high alcohol consumption are headache, nausea and vomiting, and to the extreme end cardiovascular disease, cancer at several sites, and alcoholism; these are influenced by individual variations in rate of alcohol metabolism, including its breakdown and elimination (14-17). Research suggests that genetics may determine an individual’s vulnerability to alcohol’s toxic and carcinogenic effects. Many of these unpleasant side effects are experienced due to excess in acetaldehyde production (17).

#### 2.1.3.1 Literature review between alcohol consumption and cancer

Almost 100 years ago, Lamu noticed an increase in incidence of oesophageal cancer in those who drank absinthe, a distilled, highly alcoholic (45%-75%) beverage (18). More recently, excessive alcohol consumption has been identified as a risk factor for cancer of the upper aero-digestive tract, and liver, and a probable risk factor for colorectal and breast cancer (9,19). Studies indicate a dose-dependent relationship between alcohol consumption and risk of cancer at several sites (20).

One pooled study that compiled data from 8 cohort studies in 5 countries observed a protective effect of moderate alcohol consumption on risk of colorectal cancer (21). This
study noted that the pattern of association was J-shaped with a slight protective effect when alcohol consumption fell in the range of 0 to 15 g/day compared to nondrinkers (21). However, the authors hypothesized that this pattern may have transpired as a result of collapsing both past drinkers and never-drinkers together where past drinkers may have had an elevated risk of colorectal cancer.

2.1.3.2 Possible mechanisms of alcohol-mediated carcinogenesis

The exact mechanism of alcohol-associated mutagenesis and carcinogenesis is not clear. Several mechanisms are suspected to contribute to alcohol’s toxicity.

Substantial evidence has suggested that alcohol’s toxicity could be induced by acetaldehyde, an intermediate generated during alcohol metabolism (17). Acetaldehyde is confirmed to be highly toxic, mutagenic and carcinogenic. Evidence supports that acetaldehyde rather than alcohol is responsible for the carcinogenicity linked to alcohol intake, although alcohol was suggested to facilitate the uptake of carcinogenic compounds by acting as a solvent leading them into the mucosa (3,22,23). Acetaldehyde interferes with DNA synthesis and repair and causes chromosomal rearrangements, chromatid exchanges, and point mutations in several genes (24,25). Excessive alcohol consumption has shown to alter DNA methylation in the colon of both humans and rats, which is assumed to occur through acetaldehyde (26). In rats, inhalation of this toxic metabolite prompted nasopharyngeal and laryngeal carcinoma (27). In vitro, acetaldehyde delayed cell cycle progression and enhanced cell injury associated with hyper-regeneration (23). It is well-established that acetaldehyde can impair nutrients such as folate (further discussed below) (26). As a result, the
International Agency for Research on Cancer (IARC) asserts that there is sufficient evidence to identify acetaldehyde as an established carcinogen in animals and as possibly being carcinogenic to humans (IARC Group 2B) (28).

Alcohol-induced tissue injury can also be caused through the induction of CYP2E1 that leads to generating reactive oxygen species (ROS). Cytochrome p450s are capable of generating ROS as a consequence of alcohol metabolism, which could then deplete cell reserve of antioxidants and ultimately increasing the likelihood of oxidative injury (12,13). Evidence supports that induction of CYP2E1 resulted in enhanced hepatic injury.

2.2 Polymorphic enzymes affecting the rate of alcohol metabolism

As previously noted, genetic variants exist in several classes of the ADH enzymes, including ADH1B, ADH1C, and ALDH2 (12). Additionally, CYP2E1 has been noted to be genetically polymorphic. These polymorphic genes have shown to encode for enzymes that alter the rate of alcohol oxidation and render subjects susceptible to increase in risk of several morbidities, including cancer at several sites, cardiovascular disease, alcohol liver disease, and alcoholism (23,25).

2.2.1 ADH1C enzyme

2.2.1.1 ADH1C polymorphisms

The ADH1C protein is present in approximately 13% of all ADH proteins in liver and 53% of all ADH proteins in stomach (1). Two alleles exist for the ADH1C gene, ADH1C*1 and ADH1C*2. In vitro studies have shown that ADH1C*1 enzyme has a 2.5-fold higher rate of maximal velocity ($V_{\text{max}}$) for alcohol oxidation than ADH1C*2 (5). At
CHAPTER 2: LITERATURE REVIEW

saturated concentrations of alcohol, the ADH1C*1 protein can oxidize 90 molecules of alcohol to acetaldehyde in one minute versus the slower ADH1C*2 variant, which can oxidize only 40 molecules of alcohol during the same time (1). The ADH1C*1 allele varies from ADH1C*2 in a single nucleotide on the gene locus located on chromosome 4 (8). The polymorphic site for ADH1C is an adenosine (characterizing the ADH1C*1 allele) to guanine (characterizing the ADH1C*2 allele) nucleotide base transition in exon 8 which results in the translation of the isoleucine amino acid instead of valine at residue 349 (Ile349Val) in the enzyme (7). Individuals homozygous for ADH1C*1 and ADH1C*2 alleles are classified as fast (ADH1C*1/*1) and slow (ADH1C*2/*2) metabolizers of alcohol, respectively. Subjects who are heterozygous are those that carry one copy of each allele, ADH1C*1 and ADH1C*2, (ADH1C*1/*2 genotype) and are classified as intermediate metabolizers (29).

2.2.1.2 ADH1C genotype frequencies

A meta-analysis of head and neck cancer compiled a list of observed ADH1C genotype frequencies in different ethnic groups in controls found in a collection of studies (8). One US-based study that restricted to Caucasians observed genotype frequencies of 34.6%, 48.9%, and 16.6% for homozygous fast (ADH1C*1/*1), heterozygous (ADH1C*1/*2), and homozygous slow (ADH1C*2/*2), respectively (20). The genotype frequency of ADH1C*1/*1, ADH1C*1/*2, and ADH1C*2/*2 were 64%, 32%, 4% in an African American-based population and 95%, 5%, and <1% in an Asian-based population, respectively (12,30).
2.2.1.3 Roles of ADH1C

The ADH1C enzyme has been identified as the most evolutionary constant and ancestral form of ADH discovered to-date (31). This enzyme is involved in a variety of pathways including alcohol metabolism; retinol metabolism required for the development and maintenance of many specialized epithelial tissues (32,33); metabolism of aldehyde produced by lipid peroxidation products (31,35); metabolism of compounds derived from dopamine and norepinephrine degradation (34); and possible contribution to detoxification of formaldehyde (34).

2.2.1.4 ADH1C and cancer

Thus far, evidence on the interactions between ADH1C genotype and alcohol consumption on risk of cancer at several sites have been inconclusive. A number of studies on different sites of cancer have examined the capacity of ADH1C polymorphisms to modify the effects of alcohol intake.

2.2.1.4.1 ADH1C and head and neck cancer

Six studies investigated the association between ADH1C polymorphisms and risk of head and neck cancer (8,36-39).

A pooled analysis that merged data from seven case-control studies for a total of 1,325 cases and 1,760 controls examined the association between several genes, including ADH1C and ALDH2, and risk of head and neck cancer (8). Overall, this meta-analysis did not find a statistically significant elevation in risk. Several studies found inconsistent
CHAPTER 2: LITERATURE REVIEW

results; two studies found that $ADH1C^{*2/*2}$ is associated with higher susceptibility to smoking and drinking-related head and neck squamous cell carcinoma (38,39). On the other hand, three individual studies reported that subjects with $ADH1C^{*1/*1}$ were at higher risk of head and neck cancer, including oral cancer (36,37).

2.2.1.4.2 \textit{ADH1C} and breast cancer

Four studies were identified that investigated the associations between \textit{ADH1C} polymorphisms and risk of breast cancer (40-43).

One case-control study found that women with $ADH1C^{*1/*1}$ with a lifetime alcohol consumption of 15-30 g/day had a significantly higher risk of breast cancer compared to women with $ADH1C^{*2/*2}$ and classified as non-drinkers (OR = 2.0, 95% CI, 1.1-3.5) (40). Another study found that premenopausal women who carried $ADH1C^{*1/*1}$ and consumed 6.5 drinks/month of alcohol were at a significantly elevated risk breast cancer of almost 4-fold compared to women with moderate alcohol intake (42).

In contrast, one study did not find any statistically significant effects between \textit{ADH1C} polymorphisms and risk of breast cancer (43).

2.2.1.4.3 \textit{ADH1C} and colorectal cancer

Seven studies were identified investigating the association between \textit{ADH1C} polymorphisms and risk of colorectal cancer (17,44-49).
CHAPTER 2: LITERATURE REVIEW

One study indicated that subjects who carried the $ADH1C^*1/*1$ and fell into the highest tertile of alcohol consumption (>10 drinks/week) had a significantly higher risk of colorectal cancer than those in the lowest tertile (<1 drink/week) (OR=1.76, 95% CI, 1.00-3.11); the effect between alcohol consumption and risk of colorectal cancer was weakened in subjects who carried $ADH1C^*1/*2$ and $ADH1C^*2/*2$ genotypes (44).

Another study found that the $ADH1C$ polymorphisms had no measurable association with the risk of colorectal cancer on either overall analysis or stratified analysis with alcohol use (45).

Alternatively, one study reported a significant elevation in risk of colorectal adenoma in subjects with $ADH1C^*2/*2$ genotype and high alcohol-low folate intake compared to those with a low alcohol-high folate intake (OR=17.1, 95% CI 2.1-137) with a statistically significant interaction of p<0.01 (46). Another study observed similar results suggesting that an interaction may exist between $ADH1C$ polymorphisms and alcohol consumption, where subjects with $ADH1C^*2/*2$ genotype were at higher risk of colorectal cancer with increasing consumption of alcohol (47). Furthermore, in a more recent case-control study, the authors found a significant interaction (p<0.01) between $ADH1C$ genotypes and alcohol consumption where the $ADH1C^*2/*2$ genotype was possibly associated with a increased risk of adenomatous polyps among individuals who consumed >26 g/day of alcohol (OR = 1.95, 95% CI, 0.60-6.30), a risk that is not elevated in subjects carrying $ADH1C^*1/*1$ (OR=1.09, 95% CI, 0.38-3.08), when compared to subjects with $ADH1C^*1/*1$ and consumed 0g/day (48).
2.2.2 ADH1B enzyme

2.2.2.1 ADH1B polymorphisms

Three polymorphisms of the ADH1B enzyme exist: ADH1B*1, ADH1B*2 and ADH1B*3 (5). The polymorphisms encoded by the three ADH1B alleles differ from a single amino acid residue. However, the phenotypic states of these enzymes vary markedly in their rate of alcohol metabolism, where ADH1B*2 and ADH1B*3 are at least 300 times faster in rate of oxidation than ADH1B*2 (17). In fact, the fast variants of ADH1B have been suggested to protect against alcoholism by metabolizing alcohol to acetaldehyde very readily, causing unpleasant effects of alcohol consumption (17).

2.2.2.2 ADH1B genotype frequencies

The ADH1B gene is polymorphic in Asians and African Americans, however not in Caucasians. In one study, the ADH1B genotype frequencies in Caucasians were found to be 95% for ADH1B*1/*1 genotype, and 5% having for ADH1B*1/*2 genotype (50). Conversely, the Asian population was polymorphic with the most common variant being ADH1B*2 (50). The African population was also polymorphic with the ADH1B*1 allele being most common; this population reported the presence of the third allele, ADH1B*3 (51).

2.2.3 ALDH2 enzyme

2.2.3.1 ALDH2 polymorphisms

The ALDH2 gene contains two variants, ALDH2*1 and ALDH*2. The ALDH*2 allele is inactive, which means that subjects homozygous for this allele are unable to oxidize
acetaldehyde to acetate and those who are heterozygous (ALDH2*1/*2) do so inefficiently (25,45). As a result, subjects homozygous for the inactive allele or heterozygous are likely to experience a build-up of acetaldehyde causing a toxic reaction, including flushing, higher heart rate, and nausea.

2.2.3.2 ALDH2 genotype frequencies

Evidence has shown that ALDH2 is not polymorphic in Caucasians as this ethnic group predominantly expresses the functional ALDH2*1/*1 genotype, carrying both active alleles (52). In contrast, the ALDH2*2 allele is often observed in East Asian populations, where about 30% are heterozygous (ALDH2*1/*2) and 5-10% are homozygous for the null phenotype (ALDH2*2/*2).

Because the major ethnicity in the larger study population was Caucasian (82%), this study would require a large number of subjects, which is not feasible for this master’s thesis.

2.2.4 CYP2E1 enzyme

2.2.4.1 CYP2E1 polymorphisms

Several polymorphic loci were located within the human CYP2E1 gene of which four were identifiable via RFLP (53). The polymorphisms affect transcriptional regulation or functional activity of the enzyme (11).

Two of these are two point mutations in the 5’ flanking region of the gene that include Rsa I and Pst I restriction sites, which are specific sequences of nucleotides on the DNA
that are recognized by particular restriction enzymes (Rsa I and Pst I restriction enzymes) (54). These two polymorphic sites were found to be in complete linkage disequilibrium (LD), which refers to the non-random association between alleles at adjacent genetic loci (53). The polymorphisms give rise to c1 (Pst I site -, Rsa I site +) and c2 (Pst I site +, Rsa I site-) alleles, where c2 is associated with up to 10-fold higher rate of gene transcription, protein levels, and enzyme activity than the c1 and could result in a higher exposure of acetaldehyde and ROS to the liver (53,54).

Less information is present regarding other mutations that affect CYP2E1, including Dra I and Msp I.

2.2.4.2 CYP2E1 genotype frequencies

There are significant inter-ethnic differences in CYP2E1 polymorphisms across populations. The genotype frequency for c1/c1, c1/c2, and c2/c2 is 52%, 40%, and 8% in Taiwanese; 98%, 2%, and <1% in African Americans; 92.16%, 7.68%, and 0.16% in European Americans; and 95.06%, 4.88, and 0.06% in Caucasians (53).

Similar reasons to those listed for ALDH2 and ADH1B, the low frequency of c2 allele in the Caucasian population restricted this study from exploring this gene.

Due to the low frequency of ADH1B and CYP2E1 polymorphisms and the absence of ALDH2 polymorphisms in Caucasians, it was not feasible to examine these genes as the population recruited in the larger study was predominantly Caucasian.
2.3 Homocysteine

2.3.1 Definition and properties of homocysteine

Homocysteine is a non-protein-forming sulfur amino acid that plays a central role in methionine metabolism (Figure 2.3) (55). Methionine metabolism is regulated via the bidirectional product-precursor relationship between homocysteine and methionine (55).

![Chemical structure of homocysteine](image)

**Figure 2.3.** Chemical structure of homocysteine

Methionine is an essential amino acid which is derived primarily from the diet. It is regulated in two ways; the first is through protein synthesis and second is through synthesis of S-adenosylmethionine (SAM), by means of the methionine adenosyltransferase reaction (MAT) (Figure 2.4). In mammals, there are two critical steps to the methionine metabolism: 1) transmethylation of methionine to homocysteine whereby methionine is converted to SAM, the major intracellular methyl donor, producing S-adenosylhomocysteine (SAH), which is then hydrolysed to form homocysteine; and 2) remethylation of homocysteine to methionine using the vitamin B_{12}-dependent enzyme methionine synthase (MS), which transfers a methyl group donated from 5-methyltetrahydrofolate, a major circulating form of folate in the body, to
homocysteine, ultimately remethylating homocysteine back to methionine; this cycle is ubiquitously expressed (55). An additional metabolic opportunity for the removal of homocysteine is transsulfuration, where excess homocysteine is catabolized to form cysteine. Despite the diversity of the methyl acceptors and specificity of individual enzymes, they can all be inhibited by SAH, which is the common product of SAM. Therefore, the removal of SAH is required to regulate the cycle (55).

Every mammalian cell has the ability to metabolize methionine. Homocysteine in the blood exists in different forms with the majority (~80%) bound to albumin; the remaining consists of the unbound homocysteine amino acid, including disulfide dimer and free homocysteine (56). The liver, kidney, small intestine and pancreas take up homocysteine where this amino acid undergoes catabolism (57). Accordingly, total plasma homocysteine (tHcy) is defined as the sum of all reduced and oxidized forms of homocysteine species that are in transit from the site of production to these tissues (55, 57). Most of plasma tHcy is taken up into cells and metabolized; a small percentage (0.05%) is eliminated unchanged in the urine (58).
2.3.2 Determinants of homocysteine

A number of factors are identified as the strongest determinants of tHcy concentrations in the general population such as age, sex, smoking status, coffee consumption, vitamin B₁₂, and folate intake (59). An extensive review of scientific evidence suggests that individuals with low circulating folate or vitamin B₁₂ concentrations have higher concentrations of tHcy (59-61). An inverse relationship exists where replenishment of...
folate is required to maintain low tHcy concentrations (62-64). This is explicable as folate is an important substrate in the regulation of methionine metabolism.

One study found that drinking black tea could cause a small acute increase in tHcy (65). However, both the Framingham Offspring Cohort and Hordaland Homocysteine studies found a weak positive association between tea consumption and tHcy (59,66). Other factors that were reported to be linked to tHcy concentrations are body mass index (BMI) and physical activity (67). Additionally, one study found that ethnicity is associated with tHcy concentrations where South Asian Hindus had the highest levels of tHcy (68).

### 2.3.3 Homocysteine and adverse health outcomes

Evidence supports that elevated levels of tHcy were linked to breast, cervical, colorectal, ovarian, and pancreatic cancers (69-73). Additionally, elevated levels of tHcy were shown to be linked to other morbidities such as neural tube defects, birth defects, pregnancy complications, coronary heart disease, psychiatric disorders and cognitive impairment in the elderly (55,56,58,67,75).

#### 2.3.3.1 Homocysteine and cancer

The underlying research environment for this thesis is the context of cancer etiology. The long latency period between the initiation of carcinogenesis to the diagnosis of the disease provides one motive to identify a surrogate outcome that occurs during a critical time window years prior to diagnosis. The target group for this study includes healthy subjects between 20 and 50 years where the etiology of cancer is most pertinent. As a
result, looking at a novel intermediate biomarker to cancer permits the examination of the outcome of interest many years before tumor diagnosis.

Under normal conditions, levels of tHcy are sustained at low concentrations. An increase in tHcy concentrations reflects a disturbance in the methionine-homocysteine cycle which is implicated in cancer at several sites, including colorectal, breast, leukemia cervical, ovarian, and pancreatic cancer (55,70,71,74).

2.3.4 Biomarkers

This study utilizes tHcy as an intermediate marker in cancer etiology. It represents an early effect and an intermediate between exposure and the disease of concern.

Some advantages associated with using biomarkers include: increased sensitivity and specificity in detecting exposure to carcinogenic agent, a more precise evaluation of the interplay between genetic and environmental determinants of cancer, an earlier detection of carcinogenic effects of exposure, characterization of disease, and individual risk assessment (76). Additionally, studies that have used biomarkers as intermediate or early-outcome predictors to disease could be conducted using a smaller sample size and a shorter temporal scale than if using cancer as the end point (77).

Identifying biomarkers as good candidates for surrogate endpoints could be justified by establishing causation through the criteria formulated by Dr. Austin Bradford Hill (78). One study of colorectal cancer established temporality by nesting a case-control design in a cohort with a follow-up of 10 years; this study identified that cases who were diagnosed
with colorectal cancer at the end of the follow-up had significantly higher levels of tHcy than controls \((p = 0.04)\) (71). Coincidently, the study also found that cases had much lower values of serum folate than controls \((p<0.01)\) (71). Other studies have also observed that reduced levels of folate were associated with higher risk of colorectal cancer (44-47,74,75). It is suggested that the biological plausibility criteria may also apply as folate is an important substrate that is inversely linked tHcy levels (79). One study indicated that hyperhomocysteinaemia is an independent risk factor for cancer at several sites (70). Additional findings may strengthen the argument for biological plausibility, where folate deficiency, oxidative stress, and aberrant DNA methylation were associated to hyperhomocysteinaemia and led to carcinogenesis (70). In this research, tHcy is characterized as an adequate intermediate biomarker to the risk of early phases of carcinogenesis. Results could contribute to understanding carcinogenic mechanisms that may be attributable to alcohol consumption.

2.3.5 Desirable homocysteine levels

Several studies have suggested undesirable limits of tHcy concentrations. One study that focused on vitamin B\(_{12}\) and folate deficiency suggested that tHcy values between 5 and 15 µmol/L in fasting subjects were considered normal (69). The upper limit of normal tHcy was recommended to be 12µmol/L where its clinical relevance was tested by evaluating against the outcome of coronary heart disease (CHD) (80). A number of studies of cardiovascular disease have reached a consensus indicating that the incidence of vascular disease is raised in adults with tHcy concentrations only slightly above the normal levels of 10 µmol/L (81-83). No studies have examined tHcy cutpoints relative to risk of cancer.
2.3.6 Stability of tHcy over time

Within-person measures of tHcy have been found to be relatively consistent over months and seasons, supporting the validity of using a one-time measure of tHcy. Concentrations of tHcy were found to have little seasonal variation with a reliability coefficient (R) of 0.88 between seasonal measures of tHcy in the same individuals (84). Additionally, one study found no seasonal variation in tHcy levels as reflected in their high reliability coefficient (R = 0.97) (85). One study estimated a short-term reliability coefficient of R = 0.94 over 4 weeks in 20 subjects and a longer-term reliability coefficient of R = 0.65 for 9 subjects over 30 months (86).

A significant greater difference in tHcy concentrations between individuals rather than within individuals was found in one study which conducted repeated measures of tHcy over a one-year period; this study found the within-person standard deviation (Std Dev) approximated 1 µmol/L, whereas a significantly larger between-person standard deviation approximated 2.5 µmol/L (84). The between-person coefficient of variation was 24% whereas the within-person coefficient of variation was 9%, indicating that the latter demonstrated a smaller dispersion of the tHcy measures (84). Similar results were found where tHcy collected in samples at 3-month-intervals measured a between-person coefficient of variation of 47% of tHcy and a within-person coefficient of variation of 11% (85).
2.3.7 Biological mechanism of homocysteine-mediated carcinogenesis

Disruption in the methionine-homocysteine cycle has been postulated to lead to two potential cancer-causing mechanisms: abnormal DNA methylation and the production of reactive oxygen species (ROS) (70,87).

Bulk of focus in identifying carcinogenic mechanisms associated with the disruption of this cycle has been placed on the aberration of an epigenetic mechanisms of transcriptional control involving DNA methylation, which is essential in maintaining cellular functions (20,88-91). For over 20 years, abnormal patterns of DNA methylation have been noted to occur in cancer cells (20,90). DNA methylation is a modification of a DNA strand by the addition of a methyl (-CH\(_3\)) group at the 5' position of the cytosine pyrimidine ring. Types of aberrations in this epigenetic mechanism include DNA hypomethylation and DNA hypermethylation. DNA hypomethylation is the abnormal reduction in global DNA methylation. On the other hand, DNA hypermethylation is the abnormal increase in DNA methylation which affects gene expression and conformational configuration of DNA as well as genomic instability (91). In all types of cancers, both global (or genome-wide hypomethylation) and gene-specific hypermethylation in the gene promoter regions have been observed (20). Hypermethylation has been associated with the inactivation of virtually all pathways involved in the cancer process, such as DNA repair, cell cycle regulation, carcinogen metabolism, hormonal response, and cell adherence (20). In cancer cells, global hypomethylation has encouraged genomic instability (91). Additionally, hypomethylation has shown to be an integral part of chromatin restructuring in cancer
cells that accompanies the emergence of an altered cellular phenotype (91). Furthermore, DNA hypomethylation has been implicated in the expression of proto-oncogenes.

Several studies have examined the association between abnormal DNA methylation and disruption of the methionine-homocysteine cycle. Chronic nutritional deficiencies in essential substrates and cofactors such as folate, methionine, vitamin B₆, and/or vitamin B₁₂ can perturb the methionine-homocysteine cycle (58). In effect, evidence supports that deficiency in folate causes DNA hypomethylation prior to the development of tumors (20). The mechanism in which folate deficiency and elevated measures of tHcy may promote carcinogenesis is proposed to involve increased DNA damage and mutations in tumor suppressor genes, possibly through genomic and gene-specific DNA hypomethylation (93). DNA hypomethylation has also shown to be an early and common feature in colorectal cancer (90). DNA hypomethylation can lead to genomic instability and loss of imprinting, which are features of colorectal cancer (91). One study found that decreased levels of folate are capable of inducing DNA hypomethylation in colon cells and particularly in the region of the p53 tumor suppressor gene (94). This is advocated to be mediated through the S-adenosyl-methionine (SAM) and S-adenosyl-homocysteine (SAH) as the substrate and product of this reaction; both SAM and SAH are important metabolic indicators of the cellular methylation status where elevation in homocysteine levels indicate a parallel increase in intracellular SAH which results in a product inhibition of DNA methyltransferases causing SAH-mediated DNA hypomethylation. An in vitro study suggested that accumulation of SAH may lead to inhibition of most cellular transferases and promote H₂O₂-induced DNA damage to cell lines, which could be due to impaired DNA repair capability as a result of a uracil misincorporation (95).
Elevated tHcy can also increase risk of cancer by undergoing intracellular and extracellular autooxidation in the presence of molecular oxygen which generates an excess of reactive oxygen species (ROS) that can lead to oxidative stress (70). ROS are DNA-damaging agents that can increase the mutation rate and promote and maintain the oncogenic phenotype through endogenous attacks on DNA that produce DNA adducts that are normally found in tumors (70).

2.4 Alcohol consumption and homocysteine

2.4.1 Literature

Eighteen studies were identified to examine the relationship between alcohol consumption and tHcy (79,96-112). An alcohol paradox, also referred to as the ‘French Paradox’, exists where alcohol consumption was associated to tHcy concentration in a J-shaped manner (Figure 2.5) (96). One cross-sectional study of healthy elderly adults found a dramatic J-shaped trend between the tHcy concentrations and alcohol consumption; moderate (30-60 drinks/month) consumers of alcohol had lower concentrations of tHcy than non-drinkers, heavy consumers (>60 drinks/month) had much higher concentrations of tHcy (96). Other studies found a similar but weak J-shaped pattern between alcohol consumption and tHcy levels; subjects who did not consume alcohol had higher concentrations of tHcy than moderate consumers of alcohol (79,98-100).
Chronic alcoholism is strongly associated with hyperhomocysteinemia, a medical condition characterized by abnormally large levels of tHcy in the blood (101). One pilot study monitored concentrations of tHcy within 1-2 weeks after patient admission to the hospital and observed that the group of alcohol-dependent patients hospitalized for detoxification had significantly increased concentrations of tHcy than controls; normal concentrations of tHcy were reached after 1-2 weeks of detoxification (102).

Other studies observed that in those who consumed moderate amounts of alcohol have shown lower tHcy concentrations than nondrinkers (97,100,103). Two studies found an inverse relationship between alcohol consumption and tHcy levels in moderate range only in men but not in both men and women combined (106,107).

It was recommended by that the French paradox needs further examination as tHcy were shown to increase after moderate alcohol consumption in social drinkers (108). Two studies were found that did not observe the suspected trend either (104,105). Another study reported that tHcy levels increased after moderate consumption of red wine and
spirits, but not after moderate consumption of beer which may be explained by the presence of vitamin B₆, vitamin B₁₂, and folate in beer (110-112).

### 2.4.2 Biological effects of alcohol on methionine-homocysteine cycle

The effects of alcohol consumption on tHcy have been suspected through folate deficiency as evidence suggested that serum folate levels declined within hours of alcohol ingestion (113). Increased consumption of alcohol interferes with the bioavailability of folate by diminishing its intestinal absorption, reducing its uptake and storage in the liver, inhibiting methionine synthase activity, increasing urinary loss of folate, and inducing cleavage of serum folate molecule (113-115). One study has observed that excessive alcohol intake and inadequate folate status are mechanistically linked, wherein they act synergistically to promote carcinogenesis (113).

Alcohol consumption, particularly its downstream metabolite acetaldehyde, could disrupt the methionine-homocysteine cycle by interfering with a folic acid derivative; acetaldehyde has been shown to interfere with THF by its condensation with an unsubstituted THF yielding 5, 10-methylmethylenetetrahydrofolate (Figure 2.6a) (116,117).

Interestingly, an *in vitro* experiment revealed a possible mechanism where acetaldehyde, could chemically destroy folate without recovery (118). This destruction occurs when acetaldehyde cleaves 5-methyltetrahydrofolate, also known as serum folate, at the C9-N10 bond (Figure 2.6b) (118). Another study revealed that after alcohol administration, micromolar concentrations of acetaldehyde were produced in the in the colonic mucosa...
of rats accompanied with a 50% decrease in colonic folate concentrations (119). Alcohol was also shown to block the release of folate from the hepatocyte, thereby disrupting folate supply to tissues and rapidly producing defects in cell replication (120).

There is further evidence that both acute and chronic alcohol ingestion could inhibit the activity of Methionine Synthase (MS) (121,122). Although the mechanism by which it does so remains unclear; acetaldehyde is suspected to produce adducts with the MS protein leading to its partial inactivation (Figure 2.6c) (121).

As a result of diminished folate status, inhibition of MS activity and interference with THF, the methionine-homocysteine cycle could be disrupted. This is suspected to occur in several steps. It begins with a disruption in the remethylation of homocysteine to methionine as MS could not transfer a methyl group to homocysteine (Figure 2.6d). This results in an accumulation of homocysteine (Figure 2.6e) that would then undergo autooxidation with molecular oxygen to produce an excess of ROS (Figure 2.6f). The SAM:SAH ratio could also become modified which may lead to DNA hypomethylation, a severe epigenetic abnormality (Figure 2.6g). Therefore, the pathways through which a disruption in the methionine-homocysteine cycle is theorized to lead to cancer is through DNA hypomethylation and ROS production, as previously described in section 2.3.7.
Figure 2.6 Effects of Acetaldehyde on Methionine-Homocysteine Cycle
a) Acetaldehyde undergoes condensation with THF; b) acetaldehyde cleaves serum folate at C9-N10 bond; c) acetaldehyde forms DNA adducts with MS; d) remethylation is disrupted; e) homocysteine concentrations increase; f) ROS is produced through autooxidation of homocysteine with molecular oxygen; g) SAM:SAH ratio is decreased, causing DNA hypomethylation;

2.5 *ADH1C* polymorphisms and tHcy

To date, no studies have examined the effects of *ADH1C* genotype on tHcy levels. It has been postulated that the relatively high activity of the *ADH1C*\(\ast 1\) enzyme in alcohol metabolism could manifest a potentially harmful increase in the accumulation of acetaldehyde in the liver. This could impair the overall availability of folate and MS activity that are critical for the regulation of the methionine-homocysteine cycle. As a result, individuals homozygous for *ADH1C*\(\ast 1\) are suspected to have high levels of acetaldehyde, causing deficiency in folate levels, and resulting in elevated concentrations of tHcy.
CHAPTER 2: LITERATURE REVIEW

2.6 References


38. Peters, E.S., McClean, M.D., Liu, M., Eisen, E.A., Mueller, N., and Kelsey, K.T. The ADH1C polymorphism modifies the risk of squamous cell carcinoma of the head and


CHAPTER 2: LITERATURE REVIEW


CHAPTER 3: STUDY DESIGN AND METHODS

3.1 Study design

This thesis study was nested within a larger CIHR-funded cross-sectional study led by Dr. W. King from Queen’s University serving as the Principal Investigator, Dr. T. Massey from Queen’s University, Dr. L. Dodds from Dalhousie, and S. Perkins from the University of Ottawa. The objectives of larger study are to examine the relationships between environmental factors and biomarkers of the methionine-homocysteine cycle and modification of this relationship by presence of related polymorphisms; additionally, the larger study aims to determine the association between disinfection by-products and adverse health events. Blood samples and a questionnaire were used to collect information required for the larger study, including alcohol consumption and other potential determinants of the methionine-homocysteine cycle. At the time of subject inclusion into this research, the larger study had already accrued 456 subjects along with blood samples used for genotyping and tHcy quantification.

This research is a case-control study design nested within the larger study. Information on alcohol consumption was collected using a validated global quantity-frequency (QF) questionnaire. A polymerase chain reaction restriction fragment length polymorphisms (PCR-RFLP) method was used to identify $ADH1C$ genotypes. Additional information on genotyping and alcohol quantification are further described below.
3.2 Source population and subjects

Subject recruitment for the larger study began in September 2006. The larger study employed a cross-sectional study design where data collection occurred simultaneous to subject recruitment. Subjects were healthy volunteers who were recruited through posters and pamphlets situated in Universities and hospitals in 3 areas: Kingston (ON), Ottawa (ON) and Halifax (NS); the smaller study included subjects who were recruited from Kingston (ON) and Halifax (NS).

The target population consisted of males and females ranging in age from 20-50 years. The larger study also recruited subjects with no health conditions that may be related to their measures of tHcy, including subjects with diabetes, history of angina or other vascular diseases, and pregnancy within a year prior to the index date.

3.2.1 Cases and controls

Cases were defined as those with elevated tHcy concentration $\geq 10 \, \mu\text{mol/L}$ and controls sampled from below that value. As previously discussed in section 2.3.5, several studies have shown clinical effects in adults with tHcy concentrations only slightly above the normal levels of approximately 10 $\mu$mol/L (1-3). One study transformed continuous values of tHcy into a binary outcome by using a 10 $\mu$mol/L cut-point value of tHcy as applied in this study (4). The distribution of tHcy in the larger cross-sectional study, which had complete data for 456 participants at time of subject accrual into the smaller study, demonstrated that 26% of subjects have tHcy levels equal to or greater than 10 $\mu$mol/L. As a result, this cut-point seems reasonable to use for this particular study.
Dichotomizing tHcy is valuable for several reasons. An issue that could have occurred as a result of recruiting subjects surrounding a health-conscious catchment area is that tHcy demonstrated a narrow variability; relatively few subjects were anticipated to have high levels of tHcy. This was suspected as a result of folate fortification of cereal and grain products that was initiated in 1998. Therefore, a large number of subjects would have been required for recruitment into this study in order to capture sufficient variability in tHcy to investigate the relationships of interest. This would have been unfavorable as genotyping is time consuming and costly. As a result, selecting a case-control design allowed this study to over-sample individuals with high levels of tHcy to contrast those with lower concentrations, permitting the study to obtain the maximum amount of information from the least number of subjects. This approach resulted in oversampling subjects with high tHcy values while keeping the total number of subjects for genotyping to a feasible and achievable level within the time frame for this thesis.

Cases and controls were identified with a target ratio of 1:2, respectively. All cases that met the eligibility criteria were included, and controls were frequency matched and randomly selected within each stratum of 10-year age group, sex, and centre.

3.2.2 Ethnicity

Population stratification refers to the existence of a difference in allele frequencies between cases and controls due to systematic differences in ancestry rather than association of genes with disease (5). To reduce the possibility of a false positive association between $ADH1C$ and risk of elevated tHcy due to population stratification, the population of this study was restricted to Caucasians. Evidence supported this by
noting that is ideal to restrict the study population to Caucasians when investigating the effects of $ADH1C$ polymorphisms (6).

Four genes been identified with polymorphisms that modulate the rate of alcohol metabolism and acetaldehyde production – $ADH1C$, $ADH1B$, $CYP2E1$, and $ALDH2$. The $ADH1C$ gene is highly polymorphic in Caucasians where both variants ($ADH1C*1$ and $ADH1C*2$) are expressed in almost equal frequencies. However, Asians are not polymorphic at the $ADH1C$ where the frequency of the $ADH1C*2$ occurs in less than 1% of the Asian population. Furthermore, both $ADH1B$ and $CYP2E1$ genes are not ideal for this study as both $ADH1B*2$ and $c2$ alleles encode for the fast enzyme activity of which are uncommon in Caucasians. Similarly, $ALDH2$ expresses a non-functional variant ($ALDH2*2$) that is rare in Caucasians (7).

The anticipated variation in genotype frequencies between ethnic groups in the outlined genes reflects genetically determined differences. Additionally, the ethnicity of the larger study is mainly Caucasian (82%). As a result, restricting this study to a Caucasian population permits an accurate investigation of the objectives while eliminating population stratification and the influence of other potentially interfering genes on our results, including $ALDH2$, $CYP2E1$, and $ADH1B$.

3.3 Data collection

3.3.1 Subject recruitment

Participants were recruited within the hospital adjacent to medical, research and educational institutions by a study-coordinator who is employed at each centre. The
blood draw took place in the morning after a 12-hour overnight fast and avoidance of alcohol and coffee consumption. All participants signed an informed consent prior to the blood draw followed by the completion of a questionnaire. The study coordinators met with the subject at the time of the blood draw in order to administer the questionnaire, a blood pressure assessment and to provide them with a water sample kit required to collect additional information for the larger study. Subjects were then contacted and given compensation of $30 for their participation.

3.3.2 Blood extraction and ADH1C genotyping

Blood samples were separated into different micro-tubes for biochemical and genomic analyses. Blood samples were collected from all subjects for biochemical and genomic analyses required for the larger cross-sectional study and this research study. Subject phlebotomy was scheduled for morning appointments between 8 – 10 am after a 12-hour overnight fast. This was done to ensure that recent food intake, particularly high-protein meals, coffee, and alcohol consumption, did not affect measured tHcy levels, which have been shown to fluctuate with recent food intake (8,9,10). An immediate increase in homocysteine levels was found in subjects after the end of an acute bout of high intensity exercise (9,11). As a result, blood extraction was performed 12 hours after subjects’ most recent intense exercise session.

Venipuncture was undertaken where blood was extracted from median antecubital vein and was aspirated into SST (Serum Separator Tubes) vacutainers that contain EDTA (ethylenediaminetetraacetic acid), an anticoagulant that prevents blood from clotting. Samples were immediately put on ice and centrifuged at 3300 rpm for 10 minutes at
room temperature. Serum and plasma were separated into 5 ml aliquot tubes and then stored at – 80 ºC until evaluation.

The Biochemistry laboratory of the Ottawa Hospital performed all required biochemical determinations of dietary factors including serum folate, vitamin B\textsubscript{6} and B\textsubscript{12}, Low-Density Lipoprotein (LDL) and High-Density Lipoprotein (HDL) cholesterol, triglycerides, and calcium. Total plasma homocysteine (tHcy) levels were measured using an established method in the Biochemistry laboratory under the supervision of Dr. Sherry Perkins (12).

**ADH\textsubscript{1C} Genotyping**

Laboratory procedure required for the identification of ADH\textsubscript{1C} polymorphisms was carried out at Queen’s University in Dr. Thomas Massey’s laboratory. The student researcher explored required procedures, developed, and performed PCR-RFLP for genotype identification.

DNA was first extracted from peripheral blood samples using the QIAamp DNA Blood Mini Kit (50) (Appendix D).

Identification of ADH\textsubscript{1C} genotypes was carried out using PCR-RFLP assays developed from previously published methods with minor modifications (13-15). The PCR-RFLP protocol is thoroughly outlined in the Methods section of chapter 4 under the title of “Genotype Data Collection”.
3.3.3 Measurement of alcohol consumption

From the larger study, information on alcohol consumption was collected through the study questionnaire, measured directly at the time of blood draw. The global quantity-frequency (QF) questionnaire was used to quantify alcohol consumption; it contained questions that asked participants for the frequency of alcohol consumption during the past month and the quantity of drinks consumed on those days (16). The questionnaire defined one alcoholic drink as 12oz of beer (1 entire bottle ~ 341 ml), 5oz wine (1 glass ~ 142 ml), or 1.5oz liquor (~ 43 ml), each containing more or less 0.5 oz or 14.18 g (Appendix C).

The authors of the Technical Guide recommended recoding the frequency responses based on midpoint value for the frequency of monthly consumption (16). Accordingly, values for the past month were recoded and used in the analysis as follows: “more than once a day” = 30, “about every day “= 30, “4-5 times a week” = 20, “2-3 times a week”=11, “Once a week”=4, “2-3 times a month” = 3, “once a month”=1, and “less than once a month”= 0.5 (16). These values represented the usual frequency of alcohol consumption within the past month. These numbers were then multiplied by the answer to the quantity of beverages consumed on the days they did in order to obtain the volume of alcohol consumed in standard drinks in the past month (Appendix C).

For each subject, a value for the total number of alcoholic drinks consumed within the past month was obtained by yielding the product of the quantity and frequency of alcohol consumption. The values were then converted into grams (g) where, as stated above, every beverage had approximately 14.18 g of pure alcohol (17).
3.3.4 Other variables

Variables that were related to both the exposures (alcohol and ADHIC polymorphisms) and outcome (levels of tHcy) had the potential to confound the relationships under investigation. Variables with a reported relationship with tHcy in previous literature were considered including age, sex, dietary folate, vitamin B$_{12}$, BMI, physical activity scores, smoking, and coffee and tea consumption (18). Information on age, sex, BMI, physical activity, smoking status, and coffee and tea consumption were collected through the study questionnaire at the time of blood draw. Serum folate and serum vitamin B$_{12}$ were measured in blood samples. Variables were then examined for their most effective representation into the final model.

Serum measures of folate and vitamin B$_{12}$ measured in blood samples were used as proxy measures to dietary folate and vitamin B$_{12}$, respectively. Studies have shown that serum folate, a measure of short-term intake, was strongly associated with recent use of dietary folic acid (19).

Subjects were frequency matched at the design stage on age by decade (e.g. 20-30, 30-40, 40-50), sex, and the centre from which they were recruited (e.g. Kingston or Halifax). Although these variables were adjusted for in the design stage, they were also included in the model as recommended (20).

For smoking status, subjects were categorized as non-smokers, former smokers, and current smokers. For caffeinated coffee and tea consumption variables, categories of
consumption were represented into units of cups/week (tea: 0, 0-1, >1 cup/week; coffee: 0, 0-3, >3 cups/week). Using the questionnaire, body mass index (BMI) was calculated using the ratio of weight to the square value of height in kg/m²; this was categorized according to the international classification of BMI (21). The questionnaire contained a section with the International Physical Activity Questionnaire (IPAQ) that quantified subjects’ frequency of physical activity, which were then calculated to obtain a final score of physical activity represented in units of Metabolic Equivalent (MET)-minutes-per week, as explained in the Compendium of Physical Activity (22). MET-minute scores are equivalent to kilocalories for a 60 kilogram person.

3.4 Data validity

3.4.1 ADH1C genotypes

Validity and reliability of genotyping method and results were examined. Direct sequencing of samples from different genotypes validated the PCR-RFLP method. The sequencing method was identified as the “gold standard” as it determined the exact nucleotide sequence of all sequenced samples; previously prepared PCR products and designated primers were sent in for sequencing. The procedure was carried out by Cortec DNA Service Laboratories, Inc., in Kingston Ontario. Two PCR samples of each genotype (predetermined by student investigator using PCR-RFLP) were sent for sequencing at two separate occasions.

One study recommended that at least a 10% repeat assay should be done for all subjects to ensure reliability of the genotype data (23). Over 10% repeat assay was performed for each genotype. For example, a total of 287 samples were genotyped, and of those 71
were ADH1C*1/*1 of which 10 (14%) samples were repeated, 163 were ADH1C*1/*2 of which 23 (14%) samples were repeated, and 54 were ADH1C*2/*2 of which 8 (14.8%) were repeated.

To assess validity of the genotyping data between observed and expected genotype frequencies, a Hardy-Weinberg Equilibrium (HWE) test was done in the control group. The HWE equation is \( p^2 + 2pq + q^2 = 1 \) where \( p \) and \( q \) were calculated from \( p^2 \) (observed ADH1C*1/*1 frequency), and \( q^2 \) (observed ADH1C*2/*2 frequency), respectively. The observed genotype frequencies were then compared to the expected genotype frequencies which were calculated by inputting the derived observed allele frequencies into the HWE equation (24). This served as a crude check on data, where any deviation from HWE would have indicated genotyping errors or population admixture.

### 3.4.2 Measurement of alcohol consumption

To ensure that the self-reported global QF method was accurate, one study investigated its reliability and validity for both random and systematic within-person error (25). Consistent with previous literature, the study assessed results against a daily diary intake which was used as the ‘gold standard’ as it has shown to accurately collect data on alcohol intake from subjects living in their community settings (25). Two separate administrations of the QF questionnaires, both before and after the daily record-keeping period, were highly correlated to the daily diary intake \( r = 0.95 \), suggesting adequate validity was obtained for the QF method (25). Additionally, a QF questionnaire that was administered both before and after daily diary intake had estimated a good correlation of \( r=0.93 \) suggesting that the test-retest reliability is also satisfactory. However, the study
did note that in the QF questionnaires that were administered before and after the daily
diary intake, both resulted in a slightly lower mean of alcohol consumption than the ‘gold
standard’ (25). Additional methodological issues regarding the QF questionnaire are
brought forth in Chapter 6 of this thesis.

3.5 Data management

Data on alcohol consumption and potential confounders were extracted from completed
subject questionnaires. With the help of a research assistant, data was then entered
manually into an excel spreadsheet.

Genotyping information, including PCR and RFLP, were first entered into a lab book as
per laboratory protocols; results were reviewed then transferred to the same excel
spreadsheet into which the additional data was entered.

3.6 Sample size

This study selected from a sample of 456 subjects in the larger study by December 2007.
From this, all cases (tHcy $\geq 10 \mu\text{mol/L}$) who met the inclusion criteria (having both
Caucasian parents) were included and approximately double the amount of controls was
selected randomly and frequency matched to cases by 10-year age groups, sex, and
centre.
3.7 Detectable effect

As a result of limited time and costs allotted for PCR-RFLP analysis, a total sample size of 300 was initially chosen for this study. With regard to the outlined objectives, three sets of detectable effects were expected. These objectives were examined while considering a frequency matched case:control ratio of 1:2, a two-sided type I error of 5%, and a power of 80%. Here, the detectable effects depended on the probability of exposure in controls. Calculations were completed using PS, a general program that is designed to compute sample size, detectable effect, and power using methods summarized in one study (26).

Alcohol Consumption:

The basic analysis of alcohol consumption in relation to risk of elevated tHcy is the first objective. Moderate drinkers (12.1-24.0 g/day) and heavy drinkers (>24.0 g/day) of alcohol were compared to light drinkers (≤12.0 g/day). Table 3.1 presents the parameters that were used to estimate detectable effect; the proportion of exposures was estimated from a study that dichotomized tHcy using 10 µmol/L as a cutpoint for risk of elevated tHcy (4). Approximately 45% of controls were estimated to be light drinkers, 35% moderate drinkers, and 20% heavy drinkers (4).
Table 3.1 Detectable effect calculations for alcohol consumption

<table>
<thead>
<tr>
<th>Alcohol Intake (g/day)</th>
<th>Cases</th>
<th>Controls n (%)</th>
<th>Detectable Effects (OR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤12.0</td>
<td>-</td>
<td>90 (45%)</td>
<td>Referent</td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>-</td>
<td>70 (35%)</td>
<td>0.44 and 2.17</td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>-</td>
<td>40 (20%)</td>
<td>0.34 and 2.38</td>
</tr>
<tr>
<td>Total:</td>
<td>100</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

≤12.0 g/day vs. 12.1-24.0 g/day
The contrast of the moderate consumption group against the lowest consumption group was anticipated to include approximately 160 (e.g. 70 + 90) controls and 80 (e.g. 80% of total sample multiplied by 100 cases) cases with an exposure prevalence of 44% (e.g. 70/160) in controls. The effects of association that this study can detect were calculated to be OR = 0.44 and OR = 2.17.

≤12.0 g/day vs. >24 g/day
The contrast between heavy drinkers and light drinkers was anticipated to include approximately 130 controls and 65 cases (e.g. 65% of the total sample) and an exposure prevalence of 30.8% (e.g. 40/130). The effects of association that this study can detect for these exposure groups were calculated to be OR = 0.34 and OR = 2.38.

ADH1C Polymorphisms
The second goal of this research is the basic analysis of ADH1C genotypes in relation to risk of elevated tHcy. Those in the ADH1C*1/*1 category were compared to the collapsed ADH1C*1/*2, *2/*2 category who served as the referent group. Table 3.2 presents the parameters that went into the estimates of detectable effects in this study.
The percentage of exposure was obtained from studies with outcomes including cardiovascular and cancer (15,27). The genotype frequencies were assumed to be 35% for \textit{ADH1C}*1/*1 and 65% for the collapsed \textit{ADH1C}*1/*2, *2/*2 categories.

\textbf{Table 3.2} Detectable effect calculations for \textit{ADH1C} genotypes

<table>
<thead>
<tr>
<th>\textit{ADH1C} Polymorphism</th>
<th>Cases</th>
<th>Controls n (%)</th>
<th>Detectable Effects (OR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>*1/*2, *2/*2</td>
<td>-</td>
<td>130 (65%)</td>
<td>Referent</td>
</tr>
<tr>
<td>*1/*1</td>
<td>-</td>
<td>70 (35%)</td>
<td>0.45 and 2.00</td>
</tr>
<tr>
<td>Total:</td>
<td>100</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

The exposure prevalence for those with \textit{ADH1C}*1/*1 was anticipated to be 35%. As a result, the effects of association this study can detect were calculated to be OR = 0.45 and OR = 2.00.

\textbf{Alcohol Consumption-\textit{ADH1C} Genotype Interaction}

\textbf{Table 3.3} Detectable effect calculations for \textit{ADH1C} genotypes

<table>
<thead>
<tr>
<th>Alcohol Intake (g/day)</th>
<th>\textit{ADH1C} Polymorphism</th>
<th>*1/*2, *2/*2</th>
<th>*1/*1</th>
<th>*1/*2, *2/*2</th>
<th>Detectable Effects (OR)</th>
<th>Controls (Estimated n)</th>
<th>Detectable Effects (OR)</th>
<th>Controls (Estimated n)</th>
<th>Detectable Effects (OR)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤12.0</td>
<td>*1/*2, *2/*2</td>
<td>58.5</td>
<td>Referent</td>
<td>31.5</td>
<td>0.28 and 2.81</td>
<td>90 (45%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>*1/*2, *2/*2</td>
<td>45.5</td>
<td>0.35 and 2.64</td>
<td>24.5</td>
<td>0.22 and 2.96</td>
<td>70 (35%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>*1/*2, *2/*2</td>
<td>26</td>
<td>0.23 and 2.92</td>
<td>14</td>
<td>0.06 and 3.46</td>
<td>40 (20%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>*1/*2, *2/*2</td>
<td>130 (65%)</td>
<td>70(35%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the gene-environment interaction, the above parameters were used to calculate the estimates of gene-environment association that this study can detect. Assuming the referent group to be the \textit{ADH1C}*1/*2 & *2/*2 category and ≤12.0 g/day alcohol intake...
consumption group, the detectable effects that can be calculated are outlined in the table. This objective was exploratory as the sample size required to detect the gene-environment interaction is typically large.

3.8 Design effect

Participants were recruited into the study using a method that was analogous to cluster randomization; recruitment took place from each municipality (cluster) in this study, including Kingston and Halifax. Subject recruited within each cluster were expected to be more alike than those recruited completely at random. As a result, this method of sampling may have compromised the independence between observations resulting in an underestimated estimate of the variance (27). Recently a new procedure in SAS, PROC GLIMMIX, has been developed using features from both logistic and mixed procedures. This allowed for a design effect to be considered into the model using a random effect parameter that represented centre. If this analysis resulted in a meaningful design effect, the power of this study may have been overestimated.

3.9 Data analysis

3.9.1 Preliminary analysis

For the preliminary analysis, a descriptive list of characteristics of the sample population was generated. Associations between variables and exposures were then assessed to examine potential for confounding. Crude measures of association were examined between potential confounders and risk of elevated tHcy.
Variables that were collected on a continuous scale (including serum folate, serum vitamin B₁₂, and MET-score) were examined continuously and categorically by partitioning them into tertiles or clinically meaningful cut-points. This step excluded variables that were used in the frequency-matching stage. This was done to reveal the best illustration of these variables.

Crude bivariate analyses between potential confounders, between potential confounders and the outcome, and between exposures and potential confounder were examined using Pearson Correlation for continuous variables, Chi-square tests for categorical variables, and t-tests or one-way analysis of variance (ANOVA) for continuous variables across two or more levels of categorical variables, respectively. These were done to examine potential for confounding.

Alcohol consumption was categorically divided into three groups: ≤12.0 g/day, 12.1-24.0 g/day, and >24.0 g/day. Non-drinkers and light consumers of alcohol were collapsed into one group to maximize the power required to detect an association between alcohol consumption and risk of elevated tHcy. An exploratory analysis grouped alcohol consumption into four categories to separate nondrinkers from those who consumed alcohol: 0 g/day, 0.1-12.0 g/day, 12.1-24.0 g/day, and >24.0 g/day.

The slower alcohol metabolizing ADH1C genotypes were collapsed into one group to ensure an adequate sample size for the analysis: ADH1C*1/*2 & ADH1C*2/*2 genotypes (slow metabolizers) were collapsed into one group ‘ADH1C*1/*2, *2/*2’, and ADH1C*1/*1 (fast metabolizers) genotype represented the susceptible genotype.
3.9.2 Model selection and control for confounding

This study considered potential confounders from literature that were identified as strong predictors to homocysteine. The backward elimination method was used to identify a parsimonious model predicting risk of elevated tHcy from among the potential confounders. The selection criterion applied a liberal p-value of 0.20 for inclusion into the model. This method began by including the full model adjusting for all potential confounders followed by a stepwise deletion of variables associated to the outcome beginning with the highest p-value. This continued until all variables that were predictive of the outcome with a p<0.20 were left in the model. Additional information regarding model selection is found in Chapter 4 of this thesis.

3.9.3 Main analyses

Unconditional multiple logistic regression was used to examine the relationships between alcohol consumption, ADH1C polymorphisms, and risk of elevated tHcy. All associations were interpreted as odds ratio (OR) with corresponding 95% confidence intervals (95% CI).

3.9.3.1 Effects of alcohol consumption

Deciding to examine exposure to alcohol consumption categorically rather than continuously permitted the assessment of a non-linear trend of association between alcohol consumption and risk of elevated tHcy. The ≤12.0 g/day group was specified as the referent group against which measures of association for 12.1-24.0 g/day and >24.0 g/day categories were obtained. A J-shaped relationship was anticipated where the 12.1-
24.0 g/day and >24.0 g/day groups were expected to have reduced and elevated measures of effect to the risk of elevated tHcy, respectively.

### 3.9.3.2 Effects of genotype

The main analysis that examined the measure of association of *ADH1C* polymorphisms to risk of elevated tHcy was also conducted to explore genetic susceptibility. The analysis was performed while considering the *ADH1C*\(^{*1/*1, *2/*2}\) category as the referent group.

### 3.9.3.3 Gene-environment interaction

The interaction between alcohol consumption and *ADH1C* polymorphisms to risk of elevated tHcy was explored by including a product term into the final model. Additional interactions were explored including: alcohol consumption-serum folate, *ADH1C* genotype-serum folate, and alcohol consumption-sex. These were explored by including a product term for each interaction into the logistic regression model. Sensitivity analyses were conducted to examine any potential differences in measures of association between the two method of alcohol categorization, between 4 groups of alcohol consumption (separate nondrinkers from those who consumed alcohol: 0 g/day, 0.1-12.0 g/day, 12.1-24.0 g/day, and >24.0 g/day) and 3 groups of alcohol consumption (≤ 12 g/day, 12.1-24.0 g/day, and >24.0 g/day)
3.10 Ethical considerations

To ensure that this study complied with ethical standards for research on human subjects, the original research proposal was approved by the Queen’s University Research Ethics Board. With respect to risks and benefits, informed consent and confidentiality, the following should be noted:

- There was minimal risk associated with the blood draw taken at baseline;
- The consent form affirmed that there were no direct benefit to the participants, though this study may benefit future studies;
- The participants of this study were informed of the goals of the larger study and provided their consent by signing a consent form. The objectives of this research were consistent with the original objectives described in the consent form;
- Throughout this study, there were no conditions that could have led to a loss of confidentiality or anonymity. Each blood specimen was labeled with non-identifiable subject number assigned at the start of the larger study.
3.11 References:


This manuscript was written in accordance to the instructions for authors provided by the peer-reviewed journal of *International Journal of Epidemiology*.

**Title:**

EFFECT OF ALCOHOL CONSUMPTION AND ALCOHOL DEHyDROGENASE 1C (ADH1C) POLYMORPHISMS ON TOTAL PLASMA HOMOCYSTEINE LEVELS

**Authors:**
Dana Al-Bargash¹
Will D. King¹
Thomas E. Massey²

**Affiliations:**
¹Department of Community Health and Epidemiology, Queen’s University, Kingston, Canada

²Department of Pharmacology and Toxicology, Queen’s University, Kingston, Canada
SUMMARY

BACKGROUND:
Elevated levels of total plasma homocysteine (tHcy) signify a breakdown in the methionine-homocysteine cycle. A J-shape trend between alcohol consumption and total plasma homocysteine (tHcy) has been suspected. ADH1C, an enzyme that metabolizes alcohol to acetaldehyde, may modify this relationship; ADH1C*1 has a 2.5-fold higher rate of alcohol oxidation leading to a higher accumulation of acetaldehyde than ADH1C*2. This study examined the association between alcohol consumption and risk of elevated tHcy and potential modification by polymorphisms in ADH1C.

METHODS:
A nested case-control study was conducted including 100 cases and 187 frequency matched controls ranged in age from 20-50 years. Cases were defined as having tHcy ≥ 10 µmol/L and controls were defined as those with tHcy < 10 µmol/L. Logistic regression calculated odds ratios (OR) and 95% confidence intervals (CI), adjusting for confounders.

RESULTS:
A J-shaped association was suggested between risk of elevated tHcy and alcohol consumption. Compared to ≤12.0 g/day, the adjusted OR for 12.1-24.0 g/day was 0.53 (95% CI, 0.25–1.13) and for >24.0 g/day was 1.19 (95% CI, 0.60 – 2.24). The adjusted OR for independent effect of ADH1C*1/*1 was 0.52 (95% CI, 0.27 – 1.03). The alcohol-ADH1C interaction was not statistically significant (p-value = 0.21), though among
consumers of ≤12.0 g/day, a borderline reduced measure of effect was observed in the
\textit{ADH1C}^{*1/*1} group (OR= 0.44, 95\% CI, 0.19-1.00).

CONCLUSIONS:
This research suggested a J-shaped association between of alcohol consumption and risk of elevated tHcy and a possible association between \textit{ADH1C} polymorphisms and risk of elevated tHcy.
INTRODUCTION

Total plasma homocysteine (tHcy), the total of all forms of homocysteine that exist in the plasma, has been confirmed to be an important precursor in the regulation of the methionine-homocysteine cycle (1). This thiol-containing amino acid is formed through the transmethylation of methionine (2). Malabsorption or insufficient intake of folate, vitamin B\textsubscript{12}, or vitamin B\textsubscript{6} can cause levels of tHcy to increase, indicating a breakdown in the methionine-homocysteine cycle. Higher levels of tHcy have been linked to risk of cancer at several sites, including colon and breast; a breakdown in the methionine-homocysteine cycle is suspected to cause DNA hypomethylation and oxidative stress through production of reactive oxygen species (ROS), both of which are pathways that may lead to cancer (1,3,4).

Heavy alcohol consumption has been established as a risk factor for several sites of cancer including colorectal cancer, particularly when folate levels are low (5-8). These toxic effects of heavy alcohol intake have been credited partly to the build-up of acetaldehyde following alcohol consumption. An accumulation of acetaldehyde in the blood has led authors to speculate a possible role for its involvement in the pathogenesis of tissue injury (8,9) and carcinogenesis (10). Alcohol has been identified as an antagonist of folate metabolism (11). Alcohol and acetaldehyde are suspected suppress folate bioavailability while acetaldehyde can directly destroy this substrate (6). After alcohol consumption, nearly all of the alcohol is absorbed into the blood and transported into the liver, where it is primarily metabolized (4). Over 80% of the ingested alcohol is metabolized by the enzyme alcohol dehydrogenase (ADH), generating acetaldehyde. The \textit{ADH} protein is divided into five classes based on amino acid sequence and structural
similarities. Only two ADH proteins, both class I isoenzymes, are polymorphic including \textit{ADH1B} and \textit{ADH1C}; however only \textit{ADH1C} is polymorphic in Caucasians (11). In vitro studies showed that transcription of the \textit{ADH1C*1} allele leads to an enzyme that metabolizes alcohol to acetaldehyde at a rate 2.5-fold higher than the variant \textit{(ADH1C*2)}.

A suspected J-shaped trend between alcohol consumption and risk levels of \textit{tHcy} has suggested a \textit{tHcy}-lowering mechanism during moderate consumption of alcohol (12,13). It is unclear whether the protective effect can be attributed to alcohol per se or to other components present in alcoholic beverages such as folate, vitamin B$_{12}$, or vitamin B$_{6}$ (14). On the opposite end, heavy alcohol intake has been linked to elevated levels of \textit{tHcy} (9).

This research is the first of its kind as no studies have investigated the possible modification of the relationship between alcohol consumption and risk of elevated homocysteine by any \textit{ADH} polymorphisms. \textit{ADH1C} was chosen to be the gene of interest as majority of subjects from the larger study population from which this study was nested were of Caucasian descent (80%). The trend between heavy alcohol consumption and risk of elevated \textit{tHcy} is hypothesized to be stronger in those with \textit{ADH1C*1/*1}; fast metabolizers of alcohol are expected to generate higher concentrations of acetaldehyde in the liver, followed by a decline in folate availability resulting in increasing levels of \textit{tHcy}.
METHODS

Outcome measurements

Total plasma homocysteine (tHcy) was measured in subject blood samples. Subject phlebotomy was scheduled for morning appointments between 8-10 am after a 12 hour overnight fast to ensure that recent food intakes did not affect homocysteine measures. Subjects were also required to avoid intense physical activity 12 hours prior to the blood draw. Blood was extracted from the median antecubital vein and was aspirated into vacutainers (SST and EDTA tubes). Samples were immediately put on ice and centrifuged at 3300 rpm for 10 minutes at room temperature.

This study defined cases as those with tHcy of $\geq 10$ $\mu$mol/L and controls with tHcy $< 10$ $\mu$mol/L. Nearly 26% of all subjects in the larger study (N = 456) had tHcy $\geq 10$ $\mu$mol/L.

Study population

This study recruited 287 subjects from 456 healthy participants in a larger cross-sectional study that is investigating the relationship between environmental factors and biomarkers of the methionine-homocysteine cycle. Cases and controls were sampled at 2:1 ratio and were frequency-matched on 10-year age intervals (20-30, 30-40, 40-50 years of age), sex, and centre. This study included 100 cases and 187 randomly selected controls.

The ADH1C gene can best be studied in Caucasian populations for two reasons. First, ADH1C is highly polymorphic in Caucasians, and second, other polymorphisms which are important in alcohol metabolism uncommon in Caucasians.
Eligible subjects were ranged between 20-50 years of age and born from Caucasian parents. Subjects had no health conditions that may be related to their measures of total homocysteine including pregnancy during a year prior to the index date, diabetes, history of angina, or vascular diseases. Subjects with incomplete alcohol data were excluded from the study.

A questionnaire was used to collect information required for the study, including a validated global quantity-frequency (QF) questionnaire containing 2 questions that asked participants the frequency of alcohol consumption during the past month and the quantity of drinks consumed for those days (23). This questionnaire defined one alcoholic beverage as 12 oz of beer (1 entire bottle), 5oz wine (1 glass), or 1.5 oz liquor. Information on age, sex, physical activity [indicated by Metabolic Equivalent minutes-per-week (MET)], smoking status, coffee and tea consumption, ethnicity, weight and height (where the 2 latter were used to calculate BMI) were also collected from the study questionnaire. Alcohol consumption was categorized using similar cut-points to those used in the ATTICA study, a study that investigated the effects of smoking, physical activity and dietary habits on tHcy (24).

**Genotype data collection**

Blood samples collected from the larger study were stored at -20°C. DNA was isolated using the QIamp blood kit (Qiagen, Inc.). Isolated DNA was then stored at 4°C until polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) was undertaken.
ADH1C Ile349Val polymorphisms were identified using a PCR-RFLP. Using primers described by Groppi et al. (1990), a 145-bp region of exon 8 of the ADH1C gene was amplified (24). Two primers were used where their sequences were 5'-GCTTTAAGAGTAAATATTCTGTCCC-3' and 5'-AATCTACCTCTTTCCAGAGC-3'. The PCR conditions for each sample required for amplification were 2 µl (5 pmol/µl) forward and reverse primers, 1 µl 25 mM MgCl2, 2.5 µl 10x PCR buffer, 0.2 µl (5 U/µl) DNA Taq polymerase, 0.5 µl 2.5 mM dNTP's (Biolab), and 11.7 µl water. Denaturation at 94°C for 5 minutes was followed by 32 cycles at 94°C for 1 minute, 57°C for 1 minute, and 72°C for 2 minutes, followed by 5 minutes at 72°C. To ensure that no DNA cross-contamination occurred, one sample per Mastermix of PCR reagents contained only water. Fragments were separated on a 3% agarose gel stained with ethidium bromide (1mg/ml).

RFLP was done using SSP1, a restriction enzyme that cuts double-stranded DNA following a particular specific recognition site of short nucleotide sequences (restriction sites) found in ADH1C. This fragment contained one and two SSP1 restriction sites in the wild-type (ADH1C*1) alleles and variant (ADH1C*2), respectively. Overnight (10 hours) incubation at 37°C of 10 µl of PCR produced with 2 µl of SSP1 yielded two fragments of 15 bp (base pairs) (not shown on gel) and 130 bp for homozygous wild-type (ADH1C*1/*1); four fragments of 15 bp (not shown on gel), 130 bp, 63 bp, and 67 bp for heterozygous (ADH1C*1/*2); and three fragments of 15 bp (not shown on gel), 63 bp, and 67 bp for homozygous variant (ADH1C*2/*2). These fragments were separated on a 3% agarose gel stained with ethidium bromide (1mg/ml) (Figure 4.1).
Statistical analysis

The main effects of alcohol consumption and \( ADH1C \) genotype on risk of elevated tHcy were investigated using unconditional logistic regression (25). Additional established determinants of tHcy were considered as potential confounders including: smoking, coffee consumption, tea consumption, serum folate, serum vitamin \( B_{12} \), physical activity (represented by MET minutes/week), and BMI.

The backward elimination (BE) strategy was used to create a parsimonious model adjusting for a subset of variables that were predictive for the risk of elevated tHcy. A liberal entry p-value (\( p=0.2 \)) was selected to prevent any bias that may result from underselection of important confounders (26). The final model included serum folate and vitamin \( B_{12} \), age group, sex, and centre where the latter three were forced into the model as they were used in the design stage.
RESULTS

The distributions of alcohol consumption in controls were 69% for \( \leq 12.0 \) g/day, 18% for 12.1-24.0 g/day, and 13% for \( >24.0 \) g/day, respectively. The distribution of \( ADH1C \) genotypes in controls were 73.8% for the \( ADH1C^*1/*2, *2/*2 \) group and 26.2% for the \( ADH1C^*1/*1 \) group, respectively.

The main characteristics of the study participants are shown in Table 4.1. The \( ADH1C \) genotype frequencies did not depart from the Hardy-Weinberg equilibrium among controls (\( p = 0.10 \)).

Table 4.2 outlines the adjusted measures of association between alcohol consumption and risk of elevated tHcy. After adjusting for selected covariates, the odds ratio for the 12.1-24.0 g/day category was 0.53 (95% CI, 0.25-1.13). Conversely, the \( >24.0 \) g/day group was observed to have a possible increase in odds ratio of 1.19 (95% CI, 0.60-2.34).

The second part of Table 4.2 outlines the results of \( ADH1C \) polymorphisms on risk of elevated tHcy. Although not statistically significant, the \( ADH1C^*1/*1 \) genotype had a reduced measure of effect of \( OR=0.52 \) (95% CI, 0.27-1.03) when compared to \( ADH1C^*1/*2, *2/*2 \).
### Table 4.1 Main characteristics of study population

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cases (n=100)</th>
<th>Controls (n=187)</th>
<th>Crude OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol intake, mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤12.0 g/day</td>
<td>73</td>
<td>129</td>
<td>Referent</td>
</tr>
<tr>
<td>12.1-24.0 g/day</td>
<td>10</td>
<td>33</td>
<td>0.52 (0.24 - 1.12)</td>
</tr>
<tr>
<td>&gt;24.0 g/day</td>
<td>17</td>
<td>25</td>
<td>1.17 (0.59 - 2.31)</td>
</tr>
<tr>
<td>Genotype, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADH1C*1/*2, *2/*2</td>
<td>81</td>
<td>138</td>
<td>Referent</td>
</tr>
<tr>
<td>ADH1C*1/*1</td>
<td>19</td>
<td>49</td>
<td>0.65 (0.36 - 1.17)</td>
</tr>
<tr>
<td>Age-group (yr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-29</td>
<td>40</td>
<td>74</td>
<td>Referent</td>
</tr>
<tr>
<td>30-39</td>
<td>24</td>
<td>53</td>
<td>0.87 (0.47 - 1.61)</td>
</tr>
<tr>
<td>40-50</td>
<td>36</td>
<td>60</td>
<td>1.14 (0.65 - 2.00)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>39</td>
<td>85</td>
<td>Referent</td>
</tr>
<tr>
<td>Male</td>
<td>61</td>
<td>102</td>
<td>1.29 (0.79 - 2.11)</td>
</tr>
<tr>
<td>Centre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halifax</td>
<td>55</td>
<td>101</td>
<td>Referent</td>
</tr>
<tr>
<td>Kingston</td>
<td>45</td>
<td>86</td>
<td>0.96 (0.59 - 1.57)</td>
</tr>
<tr>
<td>Smoking Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smoker</td>
<td>68</td>
<td>120</td>
<td>1.25 (0.66 - 2.34)</td>
</tr>
<tr>
<td>Former</td>
<td>18</td>
<td>39</td>
<td>Referent</td>
</tr>
<tr>
<td>Smoker</td>
<td>12</td>
<td>27</td>
<td>1.12 (0.44 - 2.48)</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below 18.5</td>
<td>4</td>
<td>5</td>
<td>1.76 (0.45-6.85)</td>
</tr>
<tr>
<td>18.5—24.9</td>
<td>45</td>
<td>101</td>
<td>Referent</td>
</tr>
<tr>
<td>25.0—29.9</td>
<td>32</td>
<td>63</td>
<td>1.15 (0.67-1.99)</td>
</tr>
<tr>
<td>30.0 and Above</td>
<td>19</td>
<td>18</td>
<td>2.32 (1.11-4.82)</td>
</tr>
<tr>
<td>Dietary Habits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coffee Intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 cups/day</td>
<td>40</td>
<td>68</td>
<td>1.35 (0.78 - 2.18)</td>
</tr>
<tr>
<td>0-3 cups/day</td>
<td>42</td>
<td>92</td>
<td>Referent</td>
</tr>
<tr>
<td>&gt;3 cups/day</td>
<td>12</td>
<td>18</td>
<td>1.46 (0.63 - 3.15)</td>
</tr>
<tr>
<td>Tea Intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 cups/day</td>
<td>47</td>
<td>81</td>
<td>1.16 (0.71 - 1.90)</td>
</tr>
<tr>
<td>0-1 cups/day</td>
<td>38</td>
<td>71</td>
<td>Referent</td>
</tr>
<tr>
<td>&gt;1 cups/day</td>
<td>15</td>
<td>35</td>
<td>0.32 (0.04 - 2.73)</td>
</tr>
<tr>
<td><strong>Mean + SD (Cases)</strong></td>
<td><strong>24.80 ± 7.76</strong></td>
<td><strong>31.26 ± 9.29</strong></td>
<td><strong>P &lt; 0.01</strong></td>
</tr>
<tr>
<td><strong>Mean + SD (Controls)</strong></td>
<td><strong>207.55 ± 118.45</strong></td>
<td><strong>245.93 ± 100.28</strong></td>
<td><strong>P &lt; 0.01</strong></td>
</tr>
<tr>
<td><strong>Unadjusted P Values</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dietary Habits</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Folic Acid (nmol/L)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vitamin B_{12} (pmol/L)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MET minutes/week</td>
<td>3426 ± 3509</td>
<td>2887 ± 2275</td>
<td>P = 0.12</td>
</tr>
</tbody>
</table>
Table 4.2 Alcohol consumption, ADH1C polymorphisms and cases of elevated tHcy

<table>
<thead>
<tr>
<th>Variables</th>
<th>Cases n=100</th>
<th>Controls n=187</th>
<th>Adjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol (g/day)$^a$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤12.0</td>
<td>73</td>
<td>129</td>
<td>Referent</td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>10</td>
<td>33</td>
<td>0.53 (0.25-1.13)</td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>17</td>
<td>25</td>
<td>1.19 (0.60-2.34)</td>
</tr>
<tr>
<td>ADH1C Polymorphism$^b$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1/*2, *2/*2</td>
<td>81</td>
<td>138</td>
<td>Referent</td>
</tr>
<tr>
<td>*1/*1</td>
<td>19</td>
<td>49</td>
<td>0.52 (0.27 - 1.03)</td>
</tr>
</tbody>
</table>

$^a$Adjusted for ADH1C, age group, sex, centre, serum folate, and serum vitamin B$_{12}$

$^b$Adjusted for alcohol intake, age group, sex, centre, serum folate, and serum vitamin B$_{12}$

There was no interaction between alcohol consumption and ADH1C (p=0.21). On the other hand, among those in the ≤12.0 g/day group, ADH1C*1/*1 had a borderline significance with OR= 0.44 (95% CI, 0.19 – 1.00). Additionally, a J-shaped trend was
observed between odds ratios of alcohol consumption and risk of elevated tHcy for each genotype group.

**DISCUSSION**

This was the first study to explore the association between alcohol consumption and tHcy while considering polymorphisms in the *ADH1C* gene.

Although results were not statistically significant, this study suggested the presence of a J-shape trend between alcohol consumption and risk of elevated tHcy (Table 4.1). In comparison to ≤12.0 g/day, results for 12.1-24.0 g/day and >24.0 g/day groups suggested reduced and increased measures of effect to the risk of elevated homocysteine, respectively.

Our results indicate that a J-shaped trend may exist between alcohol consumption and risk of elevated tHcy, a trend that has been previously observed (12,14,19). One study observed a J-shaped relationship between tHcy concentrations and amount of alcohol consumed, with a protective effect in subjects consuming ≤12.0 g/day compared to nondrinkers (14). Other studies have identified an inverse relationship between alcohol consumption in the moderate consumption range in men only and others in both men and women (19-24). The protective effect was suspected to be due to the nutritional content of beer (of folate, vitamin B\textsubscript{12} and vitamin B\textsubscript{6}) which helped maintain the tHcy levels in the normal range.

This J-shaped pattern of association is also consistent with studies of alcohol consumption and other health outcomes. The lowest risk of all-cause mortality has been
noted to be associated with moderate consumption of alcohol suggesting a protective effect (12).

No studies have investigated the gene-environment interaction between $ADH1C$ polymorphisms and alcohol consumption on tHcy levels. Recruiting a young and healthy population permits this study to investigate relationships during a meaningful time window regarding the proposed mechanisms underlying carcinogenesis. Risk of cancer at several sites, including oral, breast, and laryngeal was reported to increase in subjects carrying $ADH1C^{*1/*1}$; this genotype encodes for an enzyme with a fast rate of alcohol oxidation leading to an accumulation of acetaldehyde, which is an antagonist to folate. Folate is a critical substrate required for the regulation of tHcy levels. Therefore, it was hypothesized that a higher susceptibility to increased risk of elevated tHcy was present amongst those who carried $ADH1C^{*1/*1}$. Though results were not significant, the $ADH1C^{*1/*1}$ and $>24.0$ g/day stratum were observed to have the highest measure of effect when compared to the $ADH1C^{*1/2,*2/2}$ and $\leq 12.0$ g/day stratum (Table 4.2). However, among those who consumed $\leq 12.0$ g/day, the $ADH1C^{*1/*1}$ group was borderline significant with a reduced measure of association (OR=0.44, 95% CI, 0.19 – 1.00). Because the volume of alcohol consumed in this group is low, the reduced measure of $ADH1C^{*1/*1}$ may be due to factors that are unrelated to alcohol consumption but may be linked to the risk of elevated tHcy.

**Limitations**

Several limitations of this study should be discussed. This study had limited number of subjects which compromised the ability to identify a gene-environment interaction.
between alcohol consumption and ADH1C polymorphisms. Additionally, the small sample size limited the ability to explore the effect of non-drinkers separately from light consumers of alcohol. A sensitivity analysis was undertaken by examining nondrinkers separately from light drinkers and observed similar measures of association to those obtained here (not shown here). Another limitation is the lack of separate analysis for each genotypes; the small sample size compromised this study to detect an association as well as a gene-environment interaction between alcohol consumption and polymorphisms in ADH1C. Accordingly, a sensitivity analysis examined all genotypes separately and resulted in measures of effect to those observed in this study (not shown here).

A second limitation was the incomplete information on alcohol consumption. The questionnaire did not inquire on the type of beverage consumed, which may be critical as alcoholic beverages vary in their nutritional contents. For example, beer contains folate, vitamin B_{12} and vitamin B_{6}, all of which participate in the methionine-homocysteine cycle (19). Another study supports that the type of beverage influences tHcy, where they found that wine intake may increase tHcy concentrations, whereas beer consumption appears to have no effect on tHcy (25). Non-differential misclassification in categories of alcohol consumption was likely experienced; the true estimates of association may be underestimated.

A third limitation was the insufficient number of heavy-drinkers included in this study to detect an association. This weakened the study’s ability to examine the gene-environment interaction between alcohol consumption and ADH1C polymorphisms at heavy alcohol intake (26,27).
There could be factors that may have distorted the results obtained in this study. For example, linkage disequilibrium between \textit{ADH1C} and other gene(s), such as \textit{ADH1B}, may have altered the identified relationships. Additionally, the methylene-tetra-hydrofolate-reductase (MTHFR) gene encodes for an enzyme that is important in the methionine-homocysteine cycle where polymorphisms have shown to modify the effect of alcohol consumption on tHcy, and thus may affect results (28).

This study had limited a sample size to detect a gene-environment interaction and a strong trend between alcohol consumption and risk of elevated tHcy. Conversely, results suggest a suspected J-shaped trend between alcohol consumption and risk of elevated tHcy as well as a near significant association with \textit{ADH1C} polymorphisms. Other factors and genes that may be linked to tHcy levels and \textit{ADH1C} should be investigated to identify possible gene–gene or gene–environment interactions.

Future studies are warranted with larger sample sizes and beverage–specific measures of alcohol consumption to better investigate this suspected gene–environment interaction.
REFERENCES


5.1 Introduction

This chapter illustrates additional analyses that were undertaken with intention to complement those presented in the manuscript, including a review of the distribution of primary exposures and outcome, assessment of potential confounders, and evaluation of the robustness of the analyses to alternative representation of main variables. In addition, additional exploratory interactions were examined including alcohol–sex, alcohol–serum folate, and ADH1C–serum folate interactions in relation to risk of elevated tHcy.

This chapter also considers analyses intended to assess the methods applied in this thesis. These include consideration of folate as a confounder versus a step in the causal pathway, the potential influence of a design effect on variance, and the odds ratio as an estimate of relative risk.

Finally, validity of genotyping techniques and results are presented.

5.2 Distributions of variables

This section investigates the distribution of exposure to alcohol and tHcy

5.2.1 Distribution of alcohol

The distribution of alcohol consumption may provide some insight on the prevalence of exposure to alcohol in our sample population. This thesis grouped exposure to alcohol consumption into three categories in order to investigate a non–linear trend. A distribution of alcohol as a continuous variable was also examined.
The percentages of subjects who consumed ≤12.0 g/day, 12.1–24.0 g/day, and >24.0 g/day of alcohol in the month prior to the index date were 70.4%, 15.0%, and 14.6%. While presenting alcohol consumption on a continuous scale, the mean value was calculated to be 11.4 g/day with a standard deviation of 14.3 g/day. The median was 5.7 g/day, with an alcohol distribution that is strongly positively skewed (Figure 5.1).

This study population did not capture an adequate number of subjects with exposures to heavy alcohol consumption. This was a drawback as ADH1C polymorphisms were hypothesized to modify the effects of alcohol consumption at heavy volumes (1). One study found lower levels of tHcy existed in subjects who consumed ≤20 g/day, followed by much higher concentrations of tHcy in those consuming > 40 g/day (2). In the ATTICA study, a decline in tHcy levels was observed with alcohol intake <12 g/day compared to 24–48 g/day; highest tHcy levels were found when consumption was >48 g/day (3). In this research study, less than 10% of this study population had consumed amounts greater than 40 g/day.
5.2.2 Distribution of ADH1C Genotypes

This section examined the distribution of ADH1C genotypes and the accordance of genotyping data with the Hardy Weinberg Equilibrium Distribution.

The genotype distribution for ADH1C in the control group was in accordance with the expected Hardy–Weinberg Equilibrium distribution ($\chi^2=2.67$, df=1, p=0.10). This distribution was also compared to those in other studies. In this study, the genotype frequencies of ADH1C were as follows: ADH1C*1/*1 was 26.2%; ADH1C*1/*2 was 56.2%; and ADH1C*2/*2 was 17.7%. Similarly, one study in a sample of mainly
Caucasians found similar frequencies in genotype, where $ADHIC^{*1/*1}$ was 31.3%, $ADHIC^{*1/*2}$ was 52.5%, and $ADHIC^{*2/*2}$ was 16.4% (4). Another study based on a mixed ethnicity found genotype frequencies of 36.4% for $ADHIC^{*1/*1}$, 43.2% for $ADHIC^{*1/*2}$, and 20.3% for $ADHIC^{*2/*2}$ (5). Results indicate that the frequencies of genotypes were found to be in Hardy–Weinberg equilibrium and comparable to those found in other studies.

**5.2.3 Distribution of tHcy**

The variability of tHcy distribution in this study was examined. This could provide an indication of the relative health–status of this study population.

Data was originally collected on an approximate ratio of 1:2 frequency matching cases to controls. This study included 100 cases (35%) and 187 controls (65%) of elevated tHcy. Exploring homocysteine on a continuous scale revealed a normal distribution ranging from 3.91 µmol/L to 17.53 µmol/L, with a mean value of 8.57 µmol/L, a standard deviation of 1.95 µmol/L, a median of 8.17 µmol/L, a 25th percentile of 7.00 µmol/L, and a 75th percentile of 10.00 µmol/L (Figure 5.2).

This distribution is not expected to reflect the central tendency of tHcy values in the original study population as subjects were recruited based on their outcome values. However, the range of tHcy levels in this study could be compared to those in other studies; the range was found to be relatively consistent in studies of alcohol consumption and tHcy. However the ATTICA study that identified a J–shaped trend between alcohol
consumption and tHcy reported levels of tHcy that were slightly higher than those in this study (3).

![Figure 5.2](image)

**Figure 5.2** Distribution of homocysteine

### 5.3 Assessment of potential for confounding

In populations of observational studies, those exposed and unexposed may be dissimilar in characteristics that are also associated with the outcome. As a result, the crude relationships between exposures and outcome may reflect potential confounders. The degree of confounding is determined by the relationship between variables and the outcome (e.g. covariates), and between covariates and the exposure of interest. This
section presents an analysis of associations between exposure/outcome and variables that are suspected to confound the main investigated relationships.

5.3.1 Relationships between alcohol, \textit{ADHIC} and potential confounders

The relationships between alcohol consumption, \textit{ADHIC} polymorphisms and potential confounders were examined. Bivariate relationships between alcohol consumption and potential confounders as well as relationships between \textit{ADHIC} polymorphisms and potential confounders are presented in Table 5.1.

The relationships were assessed using Pearson product coefficient between continuous variables, 2–sample t–test or One–Way Analysis of Variance (ANOVA) between continuous and categorical values, and $\chi^2$–tests between categorical values; for small cell sizes (less than 5) between categorical variables, an exact Pearson $\chi^2$–test was used. Only p–values are presented for ANOVA and $\chi^2$–tests; Pearson’s correlation and p–values are presented for Pearson product Coefficient.

Statistical significance was considered as a p–value < 0.05. Variables that were significantly related to alcohol consumption were sex, smoking status, physical activity, and coffee consumption.

No significant relationships were observed between \textit{ADHIC} polymorphisms and examined variables. This is suspected since this gene is unlikely to be associated to environmental and modifiable factors.
The observed results between alcohol consumption and potential confounders were consistent with published literature (3). These variables have potential to confound the relationship between alcohol consumption and tHcy.

### 5.3.2 Relationship between tHcy and potential confounders

To observe whether the relationships that were suspected in literature held in this study, this section examined the associations of established determinants of homocysteine against risk of elevated tHcy. Additionally, this section assessed several variables to determine the best method of their conceptualization.

Variables that were used at the design stage for frequency matching remained in the form of representation in which they were introduced: age groups (i.e. 20–29, 30–39, 40–50 years of age), centre (i.e. Kingston/Halifax) and sex (i.e. Male/Female) were all conceptualized categorically.
Table 5.1: Relationship between covariates

<table>
<thead>
<tr>
<th></th>
<th>Alcohol Intake</th>
<th>ADH1C</th>
<th>Sex</th>
<th>Centre</th>
<th>Smoker</th>
<th>BMI</th>
<th>Physical Activity</th>
<th>Coffee Intake</th>
<th>Tea Intake</th>
<th>Serum Vitamin B&lt;sub&gt;12&lt;/sub&gt;</th>
<th>Serum Folate</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol Intake</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADH1C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>P&lt;0.01</td>
<td>P=0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centre</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.28</td>
<td>P=0.33</td>
<td></td>
<td>P=0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.05</td>
<td>P=0.61</td>
<td></td>
<td>P=0.09</td>
<td>P=0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.80</td>
<td>P=0.07</td>
<td></td>
<td>P&lt;0.01</td>
<td>P&lt;0.01</td>
<td>P=0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.03</td>
<td>P=0.54</td>
<td></td>
<td>P&lt;0.01</td>
<td>P=0.66</td>
<td>P=0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coffee Intake</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.05</td>
<td>P=0.28</td>
<td></td>
<td>P=0.03</td>
<td>P=0.25</td>
<td>P=0.01</td>
<td></td>
<td></td>
<td>P=0.14</td>
<td>P=0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tea Intake</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.18</td>
<td>P=0.87</td>
<td></td>
<td>P&lt;0.01</td>
<td>P=0.41</td>
<td>P=0.55</td>
<td></td>
<td></td>
<td>P=0.28</td>
<td>P=0.51</td>
<td>P=0.89</td>
<td></td>
</tr>
<tr>
<td>Serum Vitamin B&lt;sub&gt;12&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.86</td>
<td>P=0.97</td>
<td></td>
<td>P&lt;0.01</td>
<td>P=0.04</td>
<td>P=0.83</td>
<td></td>
<td></td>
<td>P=0.23</td>
<td>r=0.03</td>
<td>P=0.63</td>
<td>P=0.96</td>
</tr>
<tr>
<td>Serum Folate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.32</td>
<td>P=0.96</td>
<td></td>
<td>P=0.57</td>
<td>P=0.18</td>
<td>P=0.61</td>
<td></td>
<td></td>
<td>P=0.10</td>
<td>r= -0.11</td>
<td>P = 0.85</td>
<td>P=0.79</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=0.27</td>
<td>P=0.42</td>
<td></td>
<td>P&lt;0.01</td>
<td>P&lt;0.01</td>
<td>P&lt;0.01</td>
<td></td>
<td></td>
<td>P=0.07</td>
<td>P&lt;0.01</td>
<td>P=0.04</td>
<td>P=0.47</td>
</tr>
</tbody>
</table>
Three continuous variables, serum folate, serum vitamin B$_{12}$, and MET, can be represented either continuously or categorically. In order to assess the best method of their representation, Table 5.2 displayed results on categorical and continuous association for each variable to risk of elevated tHcy. Results indicated that continuous and categorical representation of serum folate and continuous representation of serum vitamin B$_{12}$ were both found to be statistically significant. Significant results were anticipated as subjects with folate or vitamin B$_{12}$ deficiency in their diet have shown elevated levels of tHcy (6,7). Though not statistically significant, physical activity was better represented when MET mins/week remained as a continuous outcome. These results suggest that all three examined variables were better left as continuous variables.

**Table 5.2: Illustration of variables in continuous and categorical forms**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Continuous Representation</th>
<th>Categorical Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P–Value</td>
<td>P–Value</td>
</tr>
<tr>
<td>Folate</td>
<td>&lt;0.01*</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Vitamin B$_{12}$</td>
<td>&lt;0.01*</td>
<td>0.11</td>
</tr>
<tr>
<td>MET– mins/week</td>
<td>0.14*</td>
<td>0.78</td>
</tr>
</tbody>
</table>

* Selected representation of variables

Some of the data that was collected was done with the implication that variables were to be conceptualized categorically, including tea consumption (i.e. 0 cups/day, 0–1 cups/day, and >1 cups/day), coffee consumption (i.e. 0 cups/day, 0–3 cups/day, and >3 cups/day), BMI (i.e. <18.5, 18.5–24.9, 25–29.9, and ≥30) and smoking status (i.e. nonsmoker, current smoker, and former smoker). Among those, a significant measure of association with elevated risk of tHcy was observed only for BMI, where a BMI score ≥ 30 had a higher risk than the normal BMI range of 18.5–24.9 (crude OR = 2.32; 95% CI,
1.11–4.82). The significant relationships were consistent with those found in previous studies (6).

5.3.3 Selection of variables into the model

The strategy chosen to select for potential confounders was the backward elimination method, which is further described in section 3.9.2. Variables were selected into the “covariate model” with intent to identify the most parsimonious model for this study. Along with those forced into the model including age groups, sex, and centre, two additional variables were selected into the model: serum folate and serum vitamin $\text{B}_{12}$.

5.4 Model fit and goodness of fit

The model fit was determined using the likelihood ratio test for overall significance and Wald test. Using the likelihood ratio test, the model was significant, $G = -2\log L \text{(with coefficients)} + 2\log L \text{(w/o coefficients)} = 48.508$, with 6 degrees of freedom (df), $p<0.01$, and the Global Wald test found $p<0.01$ where the global null hypothesis $B=0$ was rejected. This illustrated a well-fitted model whereby at least one of the predictors is significantly related to risk of elevated $\text{tHcy}$.

In terms of the covariate model, only serum folate and serum vitamin $\text{B}_{12}$ were found to be significantly related to risk of elevated $\text{tHcy}$, both at $p<0.01$. The association between variables forced into the model and risk of elevated $\text{tHcy}$ were as follows: sex – $p = 0.10$; centre – $p = 0.44$; and age group – $p = 0.36$. Although these variables were not statistically significant, they remained in the model as recommended by Breslow and Day (7).
The goodness of fit of the model describes how well the model agrees with the observed data. Generally, if the model and data are not in good agreement, then the measures of associations are not very meaningful. The Hosmer and Lemeshaw Goodness–of–fit test that is commonly used for binary response model observed $\chi^2 = 6.05$ with df = 8, and $p = 0.64$. The results were not significant indicating that the observed and predicted observations are not different from each other and that the model describes the data well.

5.5 Main analysis

Three additional interactions were explored on risk of elevated tHcy: alcohol consumption–serum folate interaction, $ADHIC$ polymorphisms–serum folate interaction, and alcohol consumption–sex interaction.

5.5.2 Serum folate

Serum folate status was explored for its potential to modify the relationship between alcohol consumption and $ADHIC$ genotype. To investigate this, folate was categorized dichotomously at the median serum folate value of 28.5 nmol/L ($\geq 28.5$ nmol/L = high folate, $< 28.5$ nmol/L = low folate).

In subjects with high serum folate status, subjects who consumed 12.1–24.0g/day and >24.0g/day observed a reduced odds ratio of 0.34 (95% CI, 0.08–1.55) and 0.63 (95% CI, 0.17–2.36), respectively (Table 5.3). In subjects with low serum folate status, subjects who consumed 12.1–24.0g/day resulted in an odds ratio of 0.95 (95% CI, 0.28–3.22) and subjects who consumed >24.0g/day resulted in almost a 2–fold increase in odds ratio.
though not statistically significant, compared to those who consumed the least amount of \( \leq 12.0 \text{ g/day} \) and in the high folate status group \((\text{OR}=1.47, 95\% \ CI, 0.42–5.16)\). Although the interaction term between alcohol consumption and folate status was not statistically significant \((p=0.31)\), a trend is suggested where the risk of elevated tHcy after alcohol consumption could be stronger in those with low serum folate. Folate–alcohol interaction results indicated a higher risk of elevated tHcy in subjects with low serum folate status who consumed \( >24.0 \text{ g/day} \) of alcohol, whereas this trend was not as apparent in subjects with high folate status.

**Table 5.3** Interaction between serum folate status and alcohol Consumption

<table>
<thead>
<tr>
<th>Alcohol Consumption (g/day)</th>
<th>Folate Status</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HIGH</td>
<td>LOW</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adjusted OR (95% CI)*</td>
<td>[Cases/Controls]</td>
<td>Adjusted OR (95% CI)*</td>
</tr>
<tr>
<td>( \leq 12.0 )</td>
<td>Referent [22, 74]</td>
<td>0.77 (0.28 – 2.16) [51, 55]</td>
<td></td>
</tr>
<tr>
<td>( 12.1-24.0 )</td>
<td>0.35 (0.09 – 1.35) [4, 23]</td>
<td>0.56 (0.14 – 2.15) [6, 10]</td>
<td></td>
</tr>
<tr>
<td>( &gt;24.0 )</td>
<td>0.73 (0.23 – 2.38) [5, 17]</td>
<td>1.47 (0.42-5.16) [12, 8]</td>
<td></td>
</tr>
<tr>
<td>p-interaction</td>
<td>0.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Adjusted for ADH1C, age group, sex, centre, serum folate, and serum vitamin B\(_{12}\)

As for \( ADH1C \) genotype, among subjects with high serum folate, the interaction term between \( ADH1C \) genotype and folate status was borderline significant. However, those encoding for \( ADH1C^{*1/*1} \) were significantly associated to reduced risk of tHcy with \( \text{OR}= 0.11 \) \((95\% \ CI, 0.01-0.90)\) (Table 5.4). Overall, results indicate that among those with high folate status, the fast metabolizers show a reduced effect of risk of elevated tHcy.
### Table 5.4 Interaction between serum folate status and ADH1C polymorphisms

<table>
<thead>
<tr>
<th>ADH1C Polymorphisms</th>
<th>Folate Status</th>
<th>Adjusted OR (95% CI)* [Cases/Controls]</th>
<th>Adjusted OR (95% CI)* [Cases/Controls]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HIGH</td>
<td>Referent [29, 85]</td>
<td>Referent [52, 53]</td>
</tr>
<tr>
<td>*1/*2, *2/*2</td>
<td>*1/*1</td>
<td>0.72 (0.31 – 1.66) [17, 20]</td>
<td>0.72 (0.31 – 1.66) [17, 20]</td>
</tr>
<tr>
<td>p-interaction</td>
<td></td>
<td>0.11 (0.01 – 0.90) [2, 29]</td>
<td>0.72 (0.31 – 1.66) [17, 20]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Referent [29, 85]</td>
<td>Referent [52, 53]</td>
</tr>
</tbody>
</table>

*Adjusted for alcohol intake, age group, sex, centre, serum folate, and serum vitamin B$_{12}$.

#### 5.5.3 Sex

Several studies have investigated the effects of alcohol consumption separately for males and females (3). This is due to the differential regulation of alcohol metabolism and methionine-homocysteine cycle between the sexes; physiological differences in alcohol metabolism exist where due to smaller amount of body water in women, they are likely to have higher blood alcohol concentration (BAC) after consuming the same amount of alcohol. Women have lower activity of the ADH enzyme in the stomach, causing a larger proportion of the ingested alcohol to reach the blood (8). The interaction between sex and alcohol consumption on risk of elevated levels of tHcy was explored (Table 5.5).

Results indicated that measure of association based on sex, illustrate that J-shape trend is suspected to be stronger in females, though not statistically significant. The interaction term, however, was not statistically significant (p=0.49) and alcohol consumption will not be investigate separately based on sex.
CHAPTER 5: SUPPLEMENTARY RESULTS

Table 5.5 Interaction between sex and alcohol Consumption

<table>
<thead>
<tr>
<th>Alcohol Consumption (g/day)</th>
<th>Sex</th>
<th>Male Adjusted OR (95% CI)* [Cases/Controls]</th>
<th>Female Adjusted OR (95% CI)* [Cases/Controls]</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤12.0</td>
<td>Referent</td>
<td>Referent [39,57]</td>
<td>Referent [34, 72]</td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>0.62 (0.23 – 1.69) [9, 24]</td>
<td>0.30 (0.03 – 2.61) [1, 9]</td>
<td></td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>1.11 (0.46 – 2.69) [13, 21]</td>
<td>2.88 (0.54 – 15.50) [4, 4]</td>
<td></td>
</tr>
</tbody>
</table>

*p-interaction 0.49

*Adjusted for alcohol intake, age group, sex, centre, serum folate, and serum vitamin B\textsubscript{12}

5.6 Sensitivity analyses

5.6.1 Alcohol intake: 3 consumption groups versus 4 consumption groups

To investigate whether results of alcohol consumption were inherently different between 3 categories of alcohol consumption (0-12.0 g/day, 12.1-24.0 g/day, >24.0 g/day) and 4 categories of alcohol consumption (0 g/day, 0.1-12.0 g/day, 12.1-24.0 g/day, >24.0 g/day), a sensitivity analysis was conducted to investigate the difference in adjusted measures of association (Table 5.6). There were no differences in adjusted odds ratio and corresponding 95% CI.

Table 5.6 Sensitivity analysis on categorization of alcohol consumption

<table>
<thead>
<tr>
<th>Alcohol Consumption (g/day)</th>
<th>Cases n = 100</th>
<th>Controls n=187</th>
<th>Adjusted Odds Ratio (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤12.0</td>
<td>73</td>
<td>129</td>
<td>Referent</td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>10</td>
<td>33</td>
<td>0.53 (0.25 – 1.13)</td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>17</td>
<td>25</td>
<td>1.19 (0.60 – 2.34)</td>
</tr>
</tbody>
</table>

| Method 2                    |               |                |                               |
| 0                           | 7             | 10             | 0.97 (0.33 – 2.87)            |
| 0.1-12.0                    | 66            | 119            | Referent                      |
| 12.1-24.0                   | 10            | 33             | 0.55 (0.23 – 1.33)            |
| >24.0                       | 17            | 25             | 1.33 (0.61 – 2.90)            |

*Adjusted for ADH1C, age group, sex, centre, serum folate, and serum vitamin B\textsubscript{12}
5.6.2 ADH1C: 2 genotype groups versus 3 genotype groups

To investigate whether categorizing subjects into the three separate genotypes would result in modified measure of association, a sensitivity analysis was conducted (Table 5.7). A very small and negligible change in odds ratio was experienced.

<table>
<thead>
<tr>
<th>Table 5.7 Sensitivity analysis on categorization of ADH1C genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADH1C Polymorphism</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>Method 1</td>
</tr>
<tr>
<td>*1/*2, *2/*2</td>
</tr>
<tr>
<td>*1/*1</td>
</tr>
<tr>
<td>Method 2</td>
</tr>
<tr>
<td>*2/*2</td>
</tr>
<tr>
<td>*1/*2</td>
</tr>
<tr>
<td>*1/*1</td>
</tr>
</tbody>
</table>

*Adjusted for alcohol intake, age group, sex, centre, serum folate, and serum vitamin B_{12}.

5.6.3 Investigating tHcy as continuous outcome

Thcy was examined as a continuous outcome to examine whether trend of alcohol consumption changes and association of ADH1C genotypes would change.

Total plasma homocysteine was normally distributed in this study population (Figure 5.2). For studies with small sample sizes less that or equal to 2000, the Shapiro-Wilk statistic is used to assess normality returning a W statistic that ranges from 0 to 1. For this research the W statistic was measured to be 0.964, indicating a normal distribution for tHcy.
Table 5.8 Linear regression analysis

<table>
<thead>
<tr>
<th>Variables</th>
<th>Parameter Estimate (95% CI)*</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol Consumption (g/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>-0.312 (-0.94 – 0.32)</td>
<td>0.33</td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>0.09 (-0.54 – 0.72)</td>
<td>0.78</td>
</tr>
<tr>
<td>ADH1C polymorphism</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1/*1</td>
<td>0.09 (-0.42 – 0.59)</td>
<td>0.74</td>
</tr>
</tbody>
</table>

*Adjusted for age group, sex, centre, serum folate, and serum vitamin B₁₂

A multiple linear regression was undertaken with parameter estimates displayed in Table 5.8. The parameter estimates are interpreted as the average difference in tHcy concentration between each variable and the referent category, which are the ≤12.0 g/day and ADH1C*1/*2, *2/*2 categories. Parameter estimates for both alcohol consumption and ADH1C genotypes were not statistically significant and not different from the null, where change in estimate = 0. However the direction of effects for alcohol consumption was consistent with previous analysis of dichotomous tHcy.

5.6.4 Removing serum folate from the model

Serum folate is suspected to occur in the causal pathway between alcohol consumption and homocysteine; it may not serve as the best proxy to dietary folate. Therefore, a sensitivity analysis was done to investigate whether the measures of association could appreciably change after excluding serum folate from the model (Table 5.9). A slight difference in odds ratio was experienced when serum folate was removed from the model. The effects of moderate alcohol consumption (12.1-24.0 g/day) resulted in a
significantly reduced measure of association with risk of elevated tHcy. The suspected J-shaped trend in alcohol consumption with risk of elevated tHcy was still present.

Table 5.9 Sensitivity analysis after removing serum folate from model

<table>
<thead>
<tr>
<th></th>
<th>Cases (n = 100)</th>
<th>Controls (n=187)</th>
<th>Adjusted Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol Consumption (g/day)³</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤12.0</td>
<td>73</td>
<td>129</td>
<td>Referent</td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>10</td>
<td>33</td>
<td>0.42 (0.18 – 0.98)</td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>17</td>
<td>25</td>
<td>1.14 (0.55 – 2.39)</td>
</tr>
<tr>
<td>ADH1C Polymorphism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1/*2, *2/*2</td>
<td>81</td>
<td>138</td>
<td>Referent</td>
</tr>
<tr>
<td>*1/*1</td>
<td>19</td>
<td>49</td>
<td>0.64 (0.34 – 1.20)</td>
</tr>
</tbody>
</table>

³ Adjusted for ADH1C, age group, sex, centre, and serum vitamin B₁₂

5.6.5 Results of design effect

As previously noted in section 3.8, the method of sampling in the larger study was similar to cluster randomization. As a result, a random effect is inserted to investigate whether this design effect would change the measure of effect and confidence interval estimates in the main analysis (Table 5.10).

In Table 5.10, measures of association and 95% CI were practically similar in both, with or without the insertion of a random effect. As a result, observations are independent, as a random effect does not seem to exist.
Table 5.10 Sensitivity analysis including random effect

<table>
<thead>
<tr>
<th>Variables</th>
<th>OR (95% CI)</th>
<th>OR with Random Effect (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol Consumption (g/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤12.0</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>0.56 (0.23 – 1.34)^a</td>
<td>0.54 (0.23 – 1.29)^b</td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>1.36 (0.63 – 2.94)^a</td>
<td>1.30 (0.60 – 2.80)^b</td>
</tr>
<tr>
<td>ADH1C polymorphism</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1/*2, *2/*2</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>*1/*1</td>
<td>0.54 (0.27 – 1.05)^c</td>
<td>0.55 (0.28 – 1.08)^d</td>
</tr>
</tbody>
</table>

^a Adjusted for genotype, age group, sex, centre, serum folate, and serum vitamin B₁₂
^b Adjusted for genotype, age group, sex, serum folate, and serum vitamin B₁₂
^c Adjusted for alcohol intake, age group, sex, centre, serum folate, and serum vitamin B₁₂
^d Adjusted for alcohol intake, age group, sex, serum folate, and serum vitamin B₁₂

5.7 Validity of genotyping techniques

DNA sequencing is regarded as the ‘gold standard’ method for identifying genotypes. To validate ADH1C genotyping, DNA sequencing of selected PCR products were found to be consistent with results obtained from the PCR-RFLP assays.

5.8 Validity of genotyping results

For quality control, approximately 10% repeat assay was performed separately for each genotype. Samples were relabeled after their selection to ensure blinding for genotype identification. Over 10% of subjects were randomly selected from each genotype to undergo repeat assays, as described in section 3.4.1. There was a 100% concordance in the repeat assay.
5.9 Odds ratio (OR) as an estimate of relative risk (RR)

This study used odds ratios as a measure of association. An odds ratio is defined as the odds of exposure among cases to the odds of exposure among controls. Results were then translated to identify which exposure group is more likely to contain more cases.

In epidemiological studies, however, relative risk (RR) is favored over odds ratio (OR) as its interpretation is more meaningful. A relative risk translates the ratio of the probability of the event occurring in the exposed group (e.g. >24.0 g/day of alcohol consumption) versus the non-exposed group (e.g. \( \leq 12.0 \) g/day of alcohol consumption). Traditionally, odds ratio can approximate the relative risk, but under particular conditions (9), where:

1. The disease must be rare;

Or

2. The controls for the study were recruited using density sampling, where they have been identified during the period when cases were occurring and not at the end of the study;

Neither one of the above conditions were met for this study. The odds ratio may likely overestimate the measures of association. As a result, this research was limited since it could not reliably estimate the RR in the traditional sense.

As described by one study, this limitation could be overcome through the application of a simple equation (10). The authors proposed an easy approximation using a simple formula, which uses the incidence of outcome of interest in the unexposed or referent group (\( P_o \)) and the estimated odds ratio:
CHAPTER 5: SUPPLEMENTARY RESULTS

\[ RR = \frac{OR}{(1 - P_0) + (P_0 \times OR)} \]

This method can also be used to estimate the corrected confidence intervals. The validity of this correction method was examined and found that the corrected risk ratio is very close to the true risk ratio and relies entirely on the appropriateness of the logistic regression model used. The estimated relative parameter estimates are illustrated in Table 5.11. RR measures do not significantly deviate from OR’s. Although the RR measures of association approached the null, the trend of alcohol consumption and the association \textit{ADH1C} polymorphisms on risk of elevated thHcy have certainly not deviated.

<table>
<thead>
<tr>
<th>Alcohol Consumption (g/day)(^a)</th>
<th>Adjusted OR (95% CI) [Cases/Controls]</th>
<th>Adjusted Relative Risk (95% CI) [Cases/Controls]</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤12.0</td>
<td>Referent [73/129]</td>
<td>Referent [73/129]</td>
</tr>
<tr>
<td>12.1-24.0</td>
<td>0.53 (0.25 – 1.13) [10/33]</td>
<td>0.64 (0.34 – 1.08) [10/33]</td>
</tr>
<tr>
<td>&gt;24.0</td>
<td>1.19 (0.60 – 2.34) [17/25]</td>
<td>1.11 (0.67 – 1.58) [17/25]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\textit{ADH1C} Polymorphism(^b)</th>
<th>Adjusted OR (95% CI) [Cases/Controls]</th>
<th>Adjusted Relative Risk (95% CI) [Cases/Controls]</th>
</tr>
</thead>
<tbody>
<tr>
<td>*1/*2, *2/*2</td>
<td>Referent [81/138]</td>
<td>Referent [81/138]</td>
</tr>
<tr>
<td>*1/*1</td>
<td>0.52 (0.27 – 1.03) [19/49]</td>
<td>0.63 (0.37 – 1.02) [19/49]</td>
</tr>
</tbody>
</table>

\(^a\) Adjusted for \textit{ADH1C}, age group, sex, centre, serum folate, and serum vitamin B\(_{12}\)

\(^b\) Adjusted for alcohol intake, age group, sex, centre, serum folate, and serum vitamin B\(_{12}\)
5.10 References


CHAPTER 6: GENERAL DISCUSSION

The objectives of this research were to examine the relationships between alcohol consumption and risk of elevated tHcy and to evaluate the effect modification of this relationship by *ADH1C* polymorphism.

6.1 Summary of key findings

6.1.1 Main results

The main effects of alcohol consumption and *ADH1C* polymorphisms on risk of elevated tHcy as well as the gene-environment interaction were discussed in the Discussion section of chapter 4. These results suggested a J-shaped trend between alcohol consumption and risk of elevated tHcy. A near significant reduced measure of association was observed for overall effects of *ADH1C*/*1/*1.

The interaction between alcohol consumption and *ADH1C* was not statistically significant. However, the fast metabolizers (ADH1C/*1/*1) and >24.0 g/day stratum was suggested to be associated with risk of elevated tHcy compared to the ≤12.0 g/day and slow metabolizers (ADH1C/*1/*2,*2/*2) stratum, although results were not statistically significant. Additionally, in subjects consuming ≤12.0 g/day, fast metabolizers had a near significant reduction in measure of association to risk of elevated in tHcy.

6.2 Methodological considerations

6.2.1 Outcome assessment

Measurement of tHcy was performed as summarized by the American Society of Human Genetics (1). Because tHcy is objectively measured through carefully performed blood
tests, the outcome is not subject to recall or selection bias. Further precaution was taken where any blood vials sent to the testing centres that were potentially mislabelled were discarded.

One problem with this method of ascertaining tHcy in this study is that only a single blood draw was done to quantify tHcy values. This introduced the possibility of misclassification in tHcy as a variation in tHcy concentration may occur if measures were ascertained at different time of day, week, month, or season. Temporal variance was avoided to some degree by taking blood samples at the same time of day, in the morning between 8-10am. Evidence confirmed that intraindividual concentrations of homocysteine were relatively constant over at least a one month period (2,3).

6.2.2 Genotype assessment

Misclassification of genotypes was likely nonexistent in this study. The process of ADH1C genotyping was done using PCR-RFLP according to an established protocol (4,5). Sequencing demonstrated 100% concordance with PCR analysis, described in section 3.3.2.

6.2.3 Exposure assessment

An important feature when quantifying alcohol consumption is to select an instrument that measures it concisely and feasibly. A brief and validated instrument was used for this purpose.
A limitation of this global quantity-frequency questionnaire is it contained no information regarding the type of beverage consumed. Previous studies have shown that light to moderate consumption of wine was protective from all-cause mortality whereas the same was not found for beer or spirits consumption (6). However, the authors identified that results may be confounded by genetic, psychosocial, or lifestyle related factors; a larger proportion of wine drinkers were of higher social class, as indicated by their assessment of educational level (6). Also, wine intake may be associated with consumption of a presumably healthy Mediterranean diet had a protective effect against coronary heart disease and a strong protective effect against cancer (6-8). Studies on tHcy, however, indicate that wine may in fact increase tHcy concentrations. On the other hand, because beer contains folate and vitamin B₆ are factors in the homocysteine methionine-cycle, this beverage was found to cause no effect on tHcy levels (9). As a result, the reduced measure of association in moderate alcohol consumption may have several explanations, moderate consumption category may have a high proportion of beer drinkers which contain dietary nutrients that are protective against high tHcy, or a high proportion of wine drinkers that may lead healthier lifestyles as a result of having certain characteristics, such as consuming a Mediterranean diet or high socio-economic status; neither was controlled for in this study.

Another limitation presented in this study is that recall on alcohol consumption was done 30 days prior to the index date. Studies recommend looking at alcohol consumption from a year prior to the index date, as it is a better representative of earlier events that might’ve encouraged the outcome of interest.
As previously discussed in the Discussion (under Limitations) section in chapter 4, information on the patterns of alcohol consumption, including binge drinking, was not captured. For example, the global QF questionnaire assumed that subjects who consume 10 alcoholic beverages in one day have an exposure equal to alcohol consumption to those who consume one drink per day over the course of 10 days. In a large prospective cohort study of more than 80,000 women, the investigators observed that very low dietary folate intake (<180 µg per day) was 2.5-fold more common among women who drank regularly (10).

Therefore by only measuring quantity and frequency of alcohol, there was no indication of the variability of a respondent’s drinking pattern. The global QF tends to underreport both frequent intake of small amounts and infrequent intakes of high amounts, preventing this study to identify binge drinkers (11). Compared with daily diary reports and other measures of alcohol, global QF tends to misclassify the amount of alcohol consumed. One report in subjects aged 23-51 who kept diaries covering 16 days observed that 31% of heavy drinkers based on a daily diary of intake were misclassified as being moderate drinkers in the QF questionnaire (12). In this study, questionnaires were completed prior to identifying the outcome. As a result, misclassification of alcohol consumption, if any, would have occurred non-differentially and would have resulted in similar patterns of misclassification in both cases and controls, causing the measures of effects to approach the null.

Conveniently, by incorporating an alcohol-consumption questionnaire into a food frequency and general questionnaire, as done with this thesis study, results were expected
to yield higher and more accurate self-report on alcohol consumption as it is less stigmatized (13). In fact, alcohol measures embedded in food frequencies showed high reproducibility, especially if compared with other food items ($r = 0.9$) (13). It was assumed that the nonstigmatized context of general food frequencies increased the reliability of alcohol consumption questionnaires.

A final limitation in exposure assessment was the low number of heavy drinkers in our sample population; in order to capture the tail end of the J-shaped trend, this study required excessive consumers of alcohol. Our study contained a small sample of heavy drinkers resulting in the inability to utilize a very high consumption category in the analysis.

### 6.2.4 Study design

The cross-sectional nature of the study design into which this smaller case-control study is nested introduces another limitation to this study. The use of this design limits the ability to identify a causal relationship, particularly regarding environmental exposure, as it is harder to establish a temporal sequence with the outcome. However, a cross-sectional design could identify an association if one exists.

Furthermore, selecting a case-control study design for this main analysis came with strengths and limitations. An advantage in dichotomizing tHcy in a population where variability in this outcome was narrow was it permitted this study to oversample subjects with high tHcy by establishing clear groups of either having the outcome or not. On the other hand, dichotomizing the outcome caused a sequential reduction in statistical power.
to detect an association (14). Also, dichotomizing tHcy had the potential for lost information and attenuated effect sizes, which essentially reduced the effectiveness of the procedure.

6.2.5 Confounders

An asset of this study is it objectively measured many strong determinants of tHcy identified in previous literature. This included unbiased biological measures of both serum folate and vitamin B\textsubscript{12} collected from blood. Additionally, the larger study included a questionnaire that collected information on lifestyle and dietary factors that have been identified as strong determinants of tHcy levels.

6.2.5.1 Serum folate

Using serum folate as a proxy to dietary folate, though accurately measured, comes with an important limitation in this study. Serum folate potentially lies in the pathway of the effect of alcohol on tHcy. Both alcohol and acetaldehyde have antagonistic effects against folate status. As a result, serum folate may act as an intermediate in the pathway that alcohol is suspected to affect. Consequently, controlling for serum folate by including it in the model could have biased the effect of alcohol consumption and interaction by ADHIC towards the null. Indeed, in a sensitivity analysis, stronger effects were observed when folate was removed from the covariate model; however it is unknown whether it occurred due to the removal of an important confounder or due to a substrate that lies in the pathway of action between exposure and outcome.
6.2.6 Sample size

The sample size for this study was predetermined to be a total of 300 subjects with a 1:2 frequency matching of cases:control. The final sample size was limited to 287 subjects with 100 cases and 187 controls. The observed size was not drastically reduced from the expected; however it reduced the ability to identify associations and a hypothesized interaction ever further. Additionally, the sample size was inadequate to explore the gene-environment interaction between alcohol consumption and $ADH1C$ polymorphisms.

6.3 Alcohol and carcinogenesis of the colon

The suggested J-shaped trend between alcohol consumption and risk of elevated tHcy was illustrated in previous study of alcohol consumption and tHcy (15,16). The reason behind the protective effect of alcohol consumption is not clear.

Consequently, since elevated tHcy was employed as an early biomarker of carcinogenesis, a similar trend is suspected to occur between alcohol consumption and risk of cancer. The mechanisms by which alcohol inflicts its carcinogenic effects are not fully understood and probably differ by target organ. However, many studies, have illustrated a dose-dependent relationship between alcohol consumption and risk of cancer in different target organs, one of which is colorectal cancer (17). A pooled analysis of primary data from 8 cohort studies in 5 countries noted that the association between alcohol consumption and colorectal cancer risk was slightly J-shaped, with a slight protective effect of alcohol consumption in range of 0 to 15 g/day (18).
In this research, the overall effect of ADH1C polymorphisms on the risk of elevated tHcy uncovered a near significant reduced measure of associations with the ADH1C*1/*1 genotype. Although a gene-environment interaction between alcohol consumption and ADH1C genotypes was not detected, it was interesting to find that the reduced measure of effect of ADH1C*1/*1 was more pronounced among those in the lowest group of alcohol consumption (≤ 12.0 g/day). Three theories were presented to explain these results:

1) The effect of alcohol consumption is modulated by ADH1C genotype

This theory is brought forth under the assumption that this study did not have sufficient subjects to identify a gene-environment interaction. It assumes that questionnaire used to collect data on alcohol consumption was likely to have under-reported the true measures of alcohol consumption. As a result, the protective association may have truly occurred as a result of alcohol consumption.

To explain the theory, the mechanism by which alcohol is metabolized must be revisited from Chapter 2. After alcohol ingestion, alcohol is rapidly absorbed through the stomach and the small intestine into the blood circulation where it reaches the liver through the portal vein to undergo a first-pass metabolism; the liver metabolizes alcohol to acetaldehyde, and ALDH2 readily metabolizes acetaldehyde to acetate. The concentration of alcohol remains constant throughout the circulation until alcohol is completely eliminated.

The rate at which the liver metabolizes alcohol depends on several factors that were previously outlined in chapter 2. One of these factors is the rate of ADH activity in the
liver, which is dependent on polymorphisms in \textit{ADH1B} and \textit{ADH1C}. For polymorphisms in \textit{ADH1C}, carrying the slow phenotype reduces ADH1C activity, which also reduces the amount of alcohol that becomes metabolized by the liver, causing alcohol to be maintained in the blood circulation longer since ADH enzymes can only metabolize an established amount of alcohol every hour (19). As a result, alcohol is distributed to different tissues, throughout the body where it would undergo metabolism.

Intestinal bacteria in the large bowel have shown ADH activity that can metabolize alcohol; however, ALDH activity is relatively low. As a result, the bacteria present in the large bowel mucosa could metabolize alcohol readily causing an increase in alcohol and acetaldehyde production and exposure to the colon epithelium that would subsequently destroy the local reservoir of folate (20). In fact, studies on rats have shown that intestinal bacteria contain high ADH activity and can oxidize alcohol producing acetaldehyde levels up to 1000-fold higher than found in the blood (21-23). Homann \textit{et al.} (2000) showed that alcohol consumption decreased levels of folate in the colonic mucus in rats by approximately 50% (23).

2. \textit{ADH1C} polymorphisms could partake in other pathways that may affect tHcy

The reduced measure of effects for \textit{ADH1C}^{*1/*1} on risk of elevated tHcy could be unrelated to alcohol metabolism and be due to other pathways in which ADH1C may partake. As discussed in chapter 2, ADH1C enzyme has been identified to metabolize other substrates, such as retinol, other aliphatic alcohols, hydroxysteroids, aldehyde, and compounds derived from dopamine and norepinephrine degradation; it may also contribute to detoxification of formaldehyde. Although limited studies have identified
association between these listed pathways and total plasma homocysteine, Future studies are warranted to investigate these pathways in relation to the methionine-homocysteine cycle.

3) ADH1C may occur in linkage disequilibrium with other genes that may affect tHcy

Linkage disequilibrium refers to the non-random association of polymorphic alleles at two or more loci, forming haplotypes, which is defined as a set of combination of polymorphisms from the two genes. Evidence suggests that ADH1C genotype is in linkage disequilibrium to several genes. One study discovered the presence of a linkage disequilibrium between dopamine receptorD2 (DRD2) and ADH1C (24). Additional evidence suggests that linkage disequilibrium also exists between ADH1B and ADH1C in Asians and possibly in Caucasians (25). Because evidence has also shown that polymorphisms in ADH1B express for enzymes that vary in their rate of alcohol oxidation, if this gene is in linkage disequilibrium with ADH1C, it could affect results by promoting false-positive or false-negative associations as the effects may be attributed to polymorphisms in ADH1B rather than ADH1C. Therefore, studies should investigate genes in linkage disequilibrium together as haplotypes rather than separately. Although this was unlikely a problem in our study due to lack of substantial evidence regarding linkage disequilibrium in Caucasians, investigation of haplotypes may warrant for future studies.

6.4 Generalizability

Since the study was restricted to a Caucasian population, the observed association between alcohol consumption, ADH1C polymorphisms, and cases of elevated tHcy may
differ in other ethnic groups. In fact, other polymorphisms involved in alcohol metabolism occur at different frequencies depending on the ethnicity; polymorphisms of \textit{ADH1B} and \textit{ALDH2} are rare in Caucasians yet common in Asians. As a result, this research study cannot be generalized to a non-Caucasian population.

6.5 Conclusion

Elevated levels of homocysteine are linked to higher risk of cancer at several sites and as well as cardiovascular disease and neural tube defects. These outcomes have also been linked to alcohol consumption and genetic susceptibility that may predispose subjects to higher risks of corresponding outcomes. This study investigated the effects of alcohol consumption on risk of elevated tHcy and while examining their interaction with \textit{ADH1C} polymorphisms. The presence of a J-shaped trend between alcohol consumption and risk of elevated tHcy assumes that a similar trend may exist with cancer. Additionally, the near significant reduced measure of association between \textit{ADH1C} polymorphisms and risk of elevated tHcy may uncover that other factors could participate in the methionine-homocysteine cycle.

6.6 Implications

Investigating a gene-environment interaction as presented in this study has important public health implications.

First, establishing a gene-environment interaction can lead to identifying a higher risk population, which can particularly benefit those who are exposed to high levels of alcohol, such as chronic alcoholics. Although alcohol may simply be a component that
CHAPTER 6: GENERAL DISCUSSION

contributes towards disease development with a weak increase in risk, its importance in research should not be dismissed as it can be of great public health importance; alcohol consumption is consumed world-wide where the ensuing effects could be severe. As a result, public health strategies can be generated for targeted intervention to subjects who are at increased exposure to alcohol and who may be genetically susceptible to its toxic effects.

Second, undertaking such studies could reveal biological mechanisms of environmental carcinogens. The majority of studies regarding relevant biological mechanisms are derived from \textit{in vitro} and animal studies. These may uncover new mechanisms of carcinogenesis; however extrapolating these results onto human models can be complex as the etiology of cancer is multi-factorial. As a result, molecular epidemiology may bridge the gap between animal studies and observational epidemiology.

\textbf{6.7 Future Direction}

Although no significant results were revealed between alcohol consumption and $ADH1C$ polymorphisms on risk of elevated tHcy, these results are not conclusive. Future direction should include studies with larger sample sizes to address the potential gene-environment interaction. Studies should also provide better quantification of alcohol consumption as beverage types may very well affect the outcome of interest. Other genes that may be in linkage with $ADH1C$, such as $ADH1B$, should be explored as their contribution to methionine metabolism could be critical.
6.7 References


APPENDIX A

Alcohol Consumption Questionnaire

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 a month

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

Conversion Factor from Question A

<table>
<thead>
<tr>
<th>Question A</th>
<th>Conversion Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>More than once a day</td>
<td>30</td>
</tr>
<tr>
<td>About every day</td>
<td>30</td>
</tr>
<tr>
<td>4 to 5 times a week</td>
<td>20</td>
</tr>
<tr>
<td>2 to 3 times a week</td>
<td>11</td>
</tr>
<tr>
<td>Once a week</td>
<td>4</td>
</tr>
<tr>
<td>2 to 3 times a month</td>
<td>3</td>
</tr>
<tr>
<td>Once a month</td>
<td>1</td>
</tr>
<tr>
<td>Less than a month</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Conversion Factor \times \text{Question B} = \text{drinks/past month}

To convert drinks of alcohol per week to grams of alcohol per day:

\text{drinks/past month} \times 0.5\text{oz/drink} \times 28.35\text{g/oz} \times \frac{1\text{ month}}{30\text{ days}} = \text{grams/day}
APPENDIX

APPENDIX B

DNA Extraction Protocol

PART ONE: Reagent Preparation
1. Qiagen Protease Stock Solution (Store @ 2-8°C or -20°C)
2. Pipet 1.2 ml protease solvent into vial containing lyophilized QIAGEN Protease indicated on the label.
3. Buffer AL (Store at room temperature)
4. Mix Buffer AL thoroughly by shaking before use.
5. Buffer AW1 (Store at room temperature), before first time use, add 25ml of Ethanol
6. Buffer AW2 (Store at room temperature), before first time use add 30 ml of ethanol

PART TWO: Blood Spin Protocol
1. Equilibrate samples to room temperature
2. Heat water bath to 56°C
3. Equilibrate Buffer AE or distilled water to room temperature
4. Check if precipitate has formed in Buffer AL (dissolve by incubating at 56°C)
5. Pipet 20µl QIAGEN Protease (Proteinase K) into the bottom of a 1.5ml microC tube
6. Add 200µl of sample to microC tube then store remaining 50ul on ice
7. Add 200µl Buffer AL to the sample
8. Mix by pulse-vortexing for 15s
9. Incubate at 56°C for 10min (Label Columns)
10. Briefly centrifuge (10s)
11. Add 200µl of ethanol (100%) to the sample
12. Mix by pulse-vortexing for 15s
13. Centrifuge briefly
14. Carefully apply the mixture to the QIAmp Spin Column (in a 2ml collection tube)
15. Centrifuge at 6000g (8000rpm) for 1min
16. Place the QIAmp Spin Column in a clean 2ml collection tube and discard the tube containing the filtrate
17. Add 500µl Buffer AW1
18. Centrifuge at 6000g (8000rpm) for 1 min.
19. Place the QIAmp Spin Column in a clean 2ml collection tube and discard the tube containing filtrate
20. Add 500µl Buffer AW2
21. Centrifuge at 20,000g (14,000rpm) for 3 min
22. Place the QIAmp Spin Column in a new 2ml collection tube and discard the tube containing the filtrate
23. Centrifuge at 20,000g (14,000rpm) for 1 min
24. Place the QIAmp Spin Column in a clean 1.5ml microC tube and discard the collection tube containing the filtrate
25. Add 200µl Buffer AE or distilled water
26. Incubate at room temperature for 5min
27. Centrifuge at 6000g (8000rpm) for 1min
28. Check DNA Concentration & Refreeze 50µl samples
29. Store @ -20°C