6-15-2021 1:00 PM

The Association of Alcohol Use and Fruit and Vegetable Consumption with Cataracts among Adults: Results from the Longitudinal Canadian National Population Health Survey (NPHS)

Yuguang Kang, *The University of Western Ontario*

Supervisor: Karp, Igor, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Epidemiology and Biostatistics

Abstract

Background: Cataracts are the leading cause of blindness globally, so advancing the understanding of their etiology is of paramount importance for the development of preventive interventions. The findings for the association of alcohol intake and fruit and vegetable consumption with cataracts in previous literature were inconsistent.

Objective: The first study objective was to assess whether alcohol intake increases the risk of cataracts among adults. The second study objective was to assess whether fruit and vegetable consumption reduces the risk of cataracts among adults.

Methods: A retrospective cohort study design was used. Data were obtained from the Household, Longitudinal component of the Canadian National Population Health Survey (NPHS) (1994-2011) cohort among adults aged 40 years or older. The first study objective used data from cycle 1 (1994/1995) through cycle 9 (2010/2011). Data for the second objective were obtained from the last five cycles of this survey (2002/2003-2010/2011). Alcohol use was measured as drinks per month. Fruit and vegetable consumption was assessed as daily servings. The subjects were followed until the occurrence of a cataract, death, end of the NPHS survey (2010/2011), or loss to follow-up, whichever came first. The research questions were addressed by fitting the Cox proportional hazards regression models with the inclusion of time-varying explanatory variables.

Results: The first study included 9,889 respondents, with 1,978 incident cataracts and an incidence rate of 19.2 per 1,000 person-years for the study population from cycle 1 to cycle 9 in NPHS. A total of 7,388 respondents who met our inclusion and exclusion criteria were identified through cycle 5 to cycle 9 for the second study, of which 1,019 developed cataracts over the follow-up period, the incidence rate was 19.7 per 1,000 person-years. After adjusting for potential confounders, the hazard ratios for alcohol intake and fruit and vegetable consumption were 1.00 (95% confidence interval [CI]: 0.95 to 1.04) and 1.04 (95% CI: 0.91 to 1.19), respectively.

Conclusion: The results suggest that alcohol use and fruit and vegetable consumption are not statistically significantly associated with the risk of cataracts.

Keywords Alcohol intake, Fruit and vegetable consumption, National Population Health Survey, Cataract, Retrospective cohort study

Summary for Lay Audience

The incidence of cataracts has grown significantly and become the leading cause of blindness globally. A cataract is a chronic condition, and many factors have been identified as being related to this condition. Risk factors such as age and sex are examples of non-modifiable factors thought to be related to developing cataracts. We assessed the association of two dietary factors (namely, alcohol use and fruit and vegetable intake) with cataracts. We used National Population Health Survey (NPHS) to examine the impact of these dietary factors on cataracts among adults. Our first study used data from NPHS to investigate whether alcohol use increases the risk of cataracts. We found that alcohol use was not statistically significantly associated with the risk of developing cataracts among adults. Our second study examined the association of fruit and vegetable intake with cataracts. We found no statistically significant association between fruit and vegetable intake and the development of cataracts.

Acknowledgements

I would like to thank my supervisor Dr Igor Karp. I am grateful for your thoughtful feedback, your unwavering support throughout this project. I would also like to thank my thesis advisory committee member: Dr Yun-Hee Choi, for her exceptional feedback for this project. This work would not have been possible without their encouragement and support emotionally and physically during this challenging time.

I would like to thank Dr Stephanie Dixon for her invaluable feedback. I would like to thank Tina, Victoria and Shane for their support during my work at RDC Western. Furthermore, I appreciate the generous help of my friends Dr Myanca Rodrigues, Dr Jason Mulimba Were, and Dr Graham D. Smith.

I would further thank all the administrative staff at the Department of Epidemiology and Biostatistics. I would also like to acknowledge many friends during my two years in our department.

Last but not least, thank you to my family. To my wife Juliana, she gave me endless support to pursue my dream. Your constant love has provided encouragement in my life. To my daughter Shimin, your smiles and words were contagious even I faced the most difficult challenges.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abbreviations

ARC- Age-Related Cataract

BMI-Body Mass Index

CAI-Computer-Assisted Interview

CHD-Coronary Heart Disease

CI-Confidence Interval

COPD- Chronic Obstructive Pulmonary Disease

DAG-Directed Acyclic Graph

DBP-Diastolic Blood Pressure

FFQ- Food Frequency Questionnaire

FV-Fruits and Vegetables

HR-Hazard Ratio

LOCS- Lens Opacities Classification System

MCMC-Markov Chain Monte Carlo

NPHS-National Population Health Survey

OR-Odds Ratio

RR-Relative Risk or Risk Ratio

ROC-Radical Oxygen Species

SBP-Systolic Blood Pressure

SD-Standard Deviation

SES-Socioeconomic Status

SMC-Swedish Mammography Cohort

TAC-Total Antioxidant Capacity

TDC-Time-Dependent Covariate

UV-Ultraviolet

Chapter 1 Introduction

This chapter provides a general introduction of cataracts. Section 1.1 summarizes the definition and classification of cataracts. Section 1.2 details the prevalence, incidence, and impact of cataracts. Section 1.3 lists important risk factors of cataracts. Section 1.4 describes the clinical presentation and diagnosis methods. Section 1.5 summarizes the primary treatment of cataracts. Section 1.6 presents the prognosis of cataracts.

1.1 Definition and Classification of Cataracts

A cataract is defined as a loss of transparency of the lens that results in visual impairment. Crystallin, a specific protein in the lens, plays an essential role to maintain the clarity of the lens. The crystalline lens contains lens fibres originated from epithelial cells. The disfunction of either crystallin or lens fibre cells or both will result in a cataract (1).

Cataracts can be classified as congenital or acquired based on the contributing causes (2). Congenital cataracts are major causes of visual loss in children (3). Our research focuses on adult cataracts, which are cataracts occurring after 20 years of age. Classification based on the predominant location of the cataracts is another common method and is more useful in the clinical setting. Nuclear, cortical and subcapsular are the three main types of cataracts (1). In most patients, two or more types of cataracts could be found at the same time.

1.2 Prevalence, Incidence, and Impact of Cataracts

1.2.1 Prevalence and Incidence of Cataracts

Several population-based studies reported the prevalence and incidence of cataracts. A study (n=1,312,051) that included only postmenopausal women in the United Kingdom reported an incidence rate of 6.4 per 1,000 person-years (4). The Salisbury Eye Evaluation study including 2,520 participants indicated African-Americans had a higher incidence than Caucasians (6.9% vs 3.0%) (5). Bastawrous et al. reported an incidence rate of 45 per 1,000 person-years for people aged 50 years or older (6). The Singapore Malay Eye Study found that cataracts were prevalent among 46% of participants aged 40 to 80 (n=3,280) (7). It also showed that the prevalence of cataracts increases dramatically with age, from 3% at 40-49 years old to 81% at 80 years and older (8). According to the National Eye Institute in the United States, it is projected the number

of people with cataracts in the US will increase to 39 million in 2030, while this number will further rise to about 50 million by 2050, which doubles the cases in 2010 (9).

To our knowledge, few studies have been conducted on the prevalence or incidence of cataracts in Canada. A report from Statistics Canada estimated that approximately 1.3 million people aged 45 years and over suffered from cataracts in 2008/2009 (10).

A recent systematic review and metal-analysis study pooled data from 45 studies with a sample size of 161,947 estimated age-standardized prevalence of cataracts globally. This study showed that the pooled prevalence of different types of cataracts is varied. The age-standardized pooled prevalence of any cataracts was estimated to be 17.2% (95% CI: 13.4 to 21.0), whereas the prevalence of cortical cataract, nuclear cataract and posterior subcapsular cataract was 8.1% (95% CI: 4.8 to 11.3), 8.2% (95% CI: 4.9 to 11.5) and 2.2% (95% CI: 1.4 to 3.1) respectively. The prevalence of cataracts increased dramatically with age. The prevalence of cataracts in individuals aged 20-39 was 3.0% (95% CI: 1.7 to 4.3), while 54.4% of people older than 60 years are estimated to have prevalent cataracts (11). The wide range of prevalence and incidence across these studies may be partly explained by the difference in the study characteristics (age, gender, race, and socioeconomic factors etc.) and methods of detection/diagnosis of the cataracts.

1.2.2 Physical Impact of Cataracts

Cataracts are the leading cause of blindness globally (12,13). It was estimated by WHO that approximately 20 million people were blind because of cataracts, which accounted for 51% of global blindness (14). Cataracts can influence daily activities, which depends upon the extent to which cataracts affect lens opacity. The Singapore Indian Eye Study showed that the impact of cataracts on daily activities is independent of refractive error, and the same study also indicated that posterior subcapsular cataracts and a combination of various cataracts have the greatest impact (15). Cataracts related to visual impairment and blindness could substantially affect health-related quality of life (16). Additionally, they could increase the risk of falls and resultant hip fractures (16,17). A recent study reported that cataracts were associated with higher mortality rates in the elderly population (18).

1.2.3 Psychological Impact of Cataracts

Functional disabilities are well-known to be associated with anxiety and depression (19,20). Several studies have been conducted to examine the association between visual impairment and depression. A longitudinal cohort study in Korea showed that both non-blindness and blindness visual impairment increase the risk of depression (21). A cohort study using data from the Canadian Longitudinal Study on Aging, including 30,097 participants, reported that the association between cataracts and three-year incidence of depression was statistically significant (relative risk [RR]:1.20, 95% CI: 1.05, 1.37) (22). A recent systematic review showed that depression was prevalent in 23% of cataract patients (23).

1.2.4 Economic Burden of Cataracts

Cataracts imposed a heavy societal, economic burden globally. It is well known that cataract surgery is a very effective procedure and can provide significant improvements in quality of life among seniors. The financial burden of cataracts primarily arises from treatment or caregiving costs and well-being losses due to visual impairment. The financial burden was estimated to be $35.4 billion (including direct medical costs, other direct costs and productivity loss) for major adult visual disorders in the United States in 2004 (24). Another study conducted in Australia also indicated that the direct financial cost of visual disorders was estimated to be $1,824.4 million in 2004, and cataracts accounted for 18% of this costs (25).

The Cost of Vision Loss in Canada, released in 2009, was the first study to use comprehensive administrative data to estimate expenditures on eye diseases resulting in visual impairment. This study showed that financial cost due to visual loss was approximately $15.6 billion in 2007, accounting for 1.19% of Canada's GDP (26).

1.3 Risk Factors for Cataracts

Many demographic factors, lifestyle factors, medications and comorbid diseases have been associated with cataracts. This section details important risk factors of cataracts.

1.3.1 Age

Age is an important non-modifiable risk factor for cataracts, as has been confirmed by numerous studies (8,27,28). The Los Angeles Latino Eye Study found that the prevalence of lens opacities increased dramatically with age. The prevalence of lens opacities was 3% in the 40 to 49-year age group, while it was tripled to 9.7% in the 50 to 59-year age group. 81.1% of people aged 80 and older had lens opacities (8). The Salisbury Eye Evaluation project, a population-based longitudinal study among 2,520 people aged 65 to 84 years old, reported that the incidence of nuclear cataract increased by 9% with a one-year increase in age (odds ratio [OR]: 1.09, 95% CI: 1.05 to 1.13) (5).

Ageing has been shown to be associated with many chronic diseases, such as coronary heart disease, diabetes, chronic obstructive pulmonary disease, and cancer. The ageing population is growing at a remarkable rate. One in six people in the world is predicted to be over 65 by 2050, up from one in eleven in 2019 (29). The frequent presence of comorbid diseases in older patients with cataracts presents a challenge to prevent and treat age-related cataracts.

1.3.2 Sex

Many studies have suggested that sex is associated with cataracts, with cataracts being more prevalent among females than males (30–34).

The effect of sex on the risk of cataracts varies with age. A cohort study conducted in Sweden indicated that the incidence of cataracts extraction in females aged over 70 years was statistically significantly higher than in males, with a relative risk (RR) being 1.71 (95% CI: 1.51 to 1.94) (31). The findings have been consistent in another population-based cohort study in Sweden, with a higher prevalence of cataract surgery for females aged between 50 to 89 (33).

The main explanations for this sex difference may be a biological influence. Oxidative DNA damage plays a vital role in the pathogenesis of cataracts (35). Estrogen can act as a mediator to play an antioxidative role by suppressing radical oxygen species (ROC) production and stimulation of antioxidant enzyme expression (36,37). The increased risk of cataracts in older women may be partly attributable to the decrease in estrogen levels at menopause. In addition,

women are more likely to seek health care than men (38). As a consequence, early-onset cataracts are more likely to be identified in female patients.

1.3.3 Race/Ethnicity

Race has been consistently indicated as a risk factor for cataracts, with disproportionate rates seen in racial and ethnic minority populations. The Salisbury Eye Evaluation study showed that African Americans have a significantly higher prevalence of cortical cataracts than Caucasians. However, a higher prevalence of posterior subcapsular cataract and nuclear cataract was also reported in Caucasians than in African Americans (39). Storey et al. also used data from the Salisbury Eye Evaluation study to examine racial differences in nuclear and cortical lens opacity (5). The study reported that African-Americans had a lower odds of nuclear opacity compared with Caucasians over a 2-years period (OR: 0.52, 95% CI: 0.35 to 0.76), while cortical opacity showed an opposite trend (OR: 1.90, 95% CI: 1.21 to 2.98). A similar pattern of associations was also presented for the progression of these two types of cataracts in this study.

A population-based study conducted in Singapore suggested that people of Asian ethnic origin had a higher age-standardized prevalence and earlier onset of cataracts than Europeans (40). Although the underlying causes of racial disparities in cataracts remain unclear, genetics might play a role in the incidence of cataracts. For example, the presence of minor allele of single-nucleotide polymorphism (SNP) rs3740030, a risk factor for age-related cataracts, has a high variance among different ethnic groups (41).

1.3.4 Socioeconomic Status

Socioeconomic status (SES) is a composite measure of income, education, occupation, and place of residence of an individual or group. SES has been shown to be a major determinant of cataracts in multiple studies. Klein et al. conducted a population-based cohort study among 4,926 participants aged 43 to 86 years old (42). The study showed that income and education were statistically significantly associated with nuclear cataracts, but not for cortical and posterior subcapsular cataracts. However, the Salisbury Eye Evaluation study showed that the cumulative incidence of either cortical or nuclear lens opacity was not statistically significantly different among those receiving education of high school or more compared with those who had lower

education (OR: 0.86, 95% CI: 0.53 to 1.40, and OR: 0.82, 95% CI: 0.57 to 1.17, respectively) (5).

1.3.5 Cigarette Smoking

Cigarette smoking has been implicated as a risk factor for various chronic illnesses (43–47), and it has also been identified as a strong risk factor for cataracts. The Physicians' Health Study with a follow-up period of 60 months reported that cigarette smoking was statistically significantly associated with cataracts after controlling other potential confounders (relative risk [RR]: 2.05, 95% Cl: 1.38 to 3.05) (48). A prospective cohort study among women aged 45 to 67 years examined the association between smoking and cataract extraction. The risk of cataract extraction was higher among those who smoked at least 65 pack-years compared with never smokers (RR: 1.63, 95% CI: 1.18 to 2.26) (49).

1.3.6 Body Mass Index

The body mass index, or BMI, which is the most widely used indicator of adiposity, is defined as the ratio of weight (in kg) divided by height^2 (in m^2). The relationship between BMI and cataracts has been investigated in several studies, but findings are not consistent.

Hiller et al. used data from the Framingham Heart Study Cohort to examine the impact of BMI on lens opacities (50). The association of BMI with cataracts varied by types of cataracts. The OR of cortical cataracts for higher BMI ($\geq$27.8 kg/m^2) was 2.19 (95% CI: 0.98 to 4.92) compared with lower BMI (<22 kg/m^2) , while the association between medium BMI (22-27.7 kg/m^2) with cortical cataract had a borderline significance (OR: 1.97, 95% CI: 1.00 to 3.89). No statistically significant association had been found for posterior subcapsular cataract and nuclear cataract.

Interestingly, different findings have been reported by the Shihpai Eye Study, which was a population-based cross-sectional study. Both nuclear and cortical cataracts were statistically significantly associated with BMI, but in the opposite direction. Higher BMI increased the odds of nuclear cataracts (OR: 1.52, 95% CI: 1.04 to 2.34), while an opposite association was observed for the cortical cataracts (OR: 0.73, 95% CI: 0.54 to 0.98). No statistically significant association for posterior subcapsular cataracts was found in this study (51).

A recent Mendelian randomization study by Tan et al. using the fat mass and obesity-related (FTO) single-nucleotide polymorphism (SNP) rs9939609 as an instrumental variable examined the association of BMI with cataracts (52).No statistically association has been found in this study after adjusting for age, sex, smoking, hypertension, and diabetes.

1.3.7 Physical Activity

Physical activity has been consistently found to be statistically significantly associated with the development of cataracts. Williams et al. reported that moderate and vigorous physical activity seemed to be protective against cataracts (53). Another population-based cohort study by Selin et al. reported that long-term physical activity was associated with a reduced risk of cataracts (54). The hazard ratio (HR) for the highest quartile of total physical activity was 0.87 (95% CI: 0.82 to 0.92) compared with the lowest quartile. Sedentary lifestyles were associated with an increased risk of cataracts. A recent meta-analysis by Jiang et al. pooled data from six prospective cohort studies, including 171,620 participants (55). Their study showed increased physical activity was inversely associated with the risk of cataracts (relative risk [RR]: 0.90, 95% CI: 0.81 to 0.99).

1.3.8 Ultraviolet Light Exposure

Ultraviolet (UV) light is one type of electromagnetic wave, and it can be categorized into three subtypes based on the wavelength: UV-A light (315-400nm), UV-B light (280-315nm), and UV-C light (100-280nm). It is well known that waves with shorter wavelengths have higher energy and are more likely to cause biological damage. Fortunately, the human cornea can act as a filter that can absorb ultraviolet radiation below 280nm, which means all UV-C is absorbed before reaching the lens (56). Thus, both UV-A and UV-B can be risk factors of cataracts (57). It has been confirmed that UV light can induce oxidation of damaged lens protein, with resultant lens opacities (58,59).

Although the impact of UV on cataracts has a solid theoretical foundation, it is hard to measure UV exposure on individuals with accuracy and precision. In the Salisbury Eye Evaluation study, UV measurement was based on the empirical model, which included a combination of an individual's ocular ambient exposure, time spent on outdoor activities and approaches of using

protective equipment. This study showed that UV-B exposure was statistically significantly associated with cortical opacity incidence (OR: 3.72, 95% CI: 1.04, 13.1) (5).

Cruickshanks et al. used data from the Beaver Dam Eye Study to examine the association between ultraviolet light exposure and lens opacities (60). The association was not statistically significant for women, while average annual ambient UVB exposure had a statistically significant association with cortical lens opacity (OR: 1.40, 95% CI: 1.06, 1.85), but not for nuclear and posterior subcapsular lens opacity. Pathologies Oculaires Liées a l'Age (POLA) reported a similar finding with the Beaver Dam Eye Study. The study suggested higher annual ambient solar radiation increased the odds of cortical and mixed cataract, but not for nuclear and posterior subcapsular cataract. After adjusting for potential confounders, the OR was 2.48 (95% CI: 1.24, 4.99) and 3.98 (95% CI: 1.98, 7.98) for cortical and mixed cataracts, respectively (61).

1.3.9 Corticosteroid Use

Corticosteroids have anti-inflammatory and immunomodulatory properties. Thus, they are widely used in many conditions, such as autoimmune diseases, asthma and chronic obstructive pulmonary disease (COPD) (62). However, long-term corticosteroid use may cause numerous side effects, and cataracts are one of them. Many studies have suggested that corticosteroid use is an important risk factor of cataracts (63–67).

A population-based case-control study using the General Practice Research Database in the United Kingdom included 15,479 patients and 15,479 controls. This study showed that long term use of corticosteroid could increase the risk of cataracts. Their association revealed a dose-response trend, with the odds ratio increasing from 0.99 (95% CI: 0.87, 1.13) with a daily dose of 400 µg to 1.69 (95% CI: 1.17, 2.43) for doses greater than 1600 µg (66).

Wang et al. conducted a longitudinal study using data from the Blue Mountains Eye Study to examine the impact of corticoid use on incident cataracts. The adjusted odds ratio of current corticoid use was 2.04 (95% CI: 1.21, 3.43) for nuclear cataract and 2.50 (95% CI: 1.33, 4.69) for posterior subcapsular cataract, showing a significant association between corticoid use and incident cataracts. However, no statistically significant association was reported for cortical cataract (65).

1.3.10 Antioxidant Vitamins/Minerals

Oxidative stress is an imbalance between the generation of free radicals and antioxidants (68). Oxidative stress is thought to play a vital role in lens opacification (69). Antioxidants are substances that can prevent the formation of free radicals (70), and thus, they were hypothesized to have a protective role in cataract formation. Dietary supplements (e.g., vitamin C, vitamin E, carotenoids and polyphenols) are important sources of exogenous antioxidants (71). Numerous studies have been conducted exploring the association between dietary supplements and cataracts. However, the findings of these existing studies have been inconsistent.

Selin et al. conducted a prospective cohort study among men to examine the association of vitamin intake with age-related cataracts. Surprisingly, the study reported that both vitamin C use (HR:1.21, 95% CI: 1.04-1.41) and vitamin E (HR: 1.56, 95% CI: 1.12 to 2.26) were associated with cataracts 1.59 (72).

Jiang et al. conducted a meta-analysis that included 8 RCTs and 12 cohort studies. Several vitamins, such as vitamin A, vitamin C, vitamin E, β-carotene and lutein or zeaxanthin, were found to have a negative association with cataracts, while the association of vitamin E or β-carotene with cataract was not statistically significant in the RCTs (73).

1.3.11 Myopia

Although several studies have been conducted to examine the association between myopia and cataracts, these studies showed inconsistent and conflicting results. The Visual Impairment Project conducted in Australia showed that myopia was statistically significantly associated with cortical cataracts (relative risk [RR]: 2.2, 95% CI: 1.4-3.4), whereas the association of myopia with posterior subcapsular and nuclear cataracts was not statistically significant (RR: 1.1, 95% CI: 0.69-1.9 and RR: 1.0, 95% CI: 0.62-1.6 respectively) (74).

However, the Blue Mountains Eye Study showed the association of myopia with cortical and nuclear cataracts was not statistically significant. The posterior subcapsular cataracts was statistically significantly associated with myopia (OR: 2.6, 95% CI: 1.4-5.0)(75).

A meta-analysis showed that myopia was only statistically significantly associated with the prevalence of nuclear and posterior subcapsular cataracts based on the pooled data from seven

cross-sectional studies and one case-control study, while myopia was not statistically significantly associated with the incidence of any types of cataracts. However, the included studies showed high heterogeneity, based on the evaluation of the I^2 statistic (76).

1.3.12 Diabetes

Diabetes has been implicated as a risk factor for cataracts. Study findings have consistently shown that diabetes is significantly and positively associated with the cataracts (77–84).

A study by Machan et al., which used the Waterloo Eye Study data and included 6,397 patients (ages 1 to 93 years) in Canada, found that type 2 diabetes was positively associated with age-related cataracts (OR: 1.60, 95% CI: 1.13-2.27) (78). In the further analysis, the authors reported that diabetes was positively associated with cortical cataracts (OR: 1.62, 95% CI: 1.14-2.29) and nuclear cataracts (OR: 1.37, 95% CI: 1.02-1.83). However, the association with posterior subcapsular cataracts was not statistically significant (OR: 1.33, 95% CI: 0.90-1.96) (78).

Interestingly, one paper by Jacques et al., which examined the association of weight status and diabetes with age-related lens opacities using the Nurses' Health Study (NHS) data (n=466, ages 53 to 73 years), found that women with diabetes were more likely to have posterior subcapsular cataracts (OR: 4.1, 95% CI: 1.8-9.4). However, the association was not statistically significant for cortical and nuclear cataracts (81).

A meta-analysis based on the eight studies reported that the odds of having any cataracts was higher among diabetic patients than those without diabetes (OR=1.97, 95% CI: 1.45-2.67). The additional analysis showed that the risk of cortical cataracts (OR=1.48, 95% CI: 1.47-1.91) and posterior subcapsular cataracts (OR=1.55, 95% CI: 1.27 to 1.90) was higher in type 2 diabetic patients, while the association was not statistically significant for nuclear cataracts (OR=1.36, 95% CI: 0.97 to 1.90) (77).

1.3.13 Hypertension

Evidence from numerous studies has shown the association between hypertension and cataracts, although the mechanism is not well understood (85–88).

Klein et al. investigated the association of hypertension and cataracts using data from the Beaver Dam Eye Study. The authors found that hypertension was statistically significantly associated

with posterior subcapsular cataracts (OR: 1.39, 95% CI: 1.05-1.84), while any cataracts, cortical and nuclear cataracts were not statistically significantly associated with hypertension (85).

Goodrich et al. used data from the Blue Mountains Eye study to examine the relationship of cardiovascular disease factors with cataracts. Interestingly, the findings of this study indicated that hypertension was negatively associated with nuclear cataracts (OR: 0.8, 95% CI: 0.6-0.9). However, the association was not statistically significant for posterior subcapsular and cortical cataracts (88).

One paper by Schaumberg et al. examined the association of hypertension with age-related cataracts using the Physicians' Health Study (PHS) data (n=22,071, ages 40 to 84 years). Systolic blood pressure (SBP) and diastolic blood pressure (DBP) was investigated separately in this study. The study showed that patients with SBP≥150mmHg were more likely to have incident nuclear cataracts (HR: 1.33, 95% CI: 1.03-1.70), but not for cortical cataracts (HR: 1.10, 95% CI: 0.75-1.62) and posterior subcapsular cataracts (HR: 1.14, 95% CI: 0.78-1.68), compared those with SBP<120mmHg. The DBP was not found to be statistically significantly associated with the risk of cataracts (86).

1.3.14 Genetics

Genetics plays an essential role in the development of cataracts. Inherited genetic variation contributes both directly and indirectly to the pathogenesis of cataracts. A congenital cataract is usually a consequence of mutations in crystallin or other lens proteins, and at least 44 genetic loci have been found to be associated with primary congenital cataract (89). Age-related cataracts are thought to be a multifactorial disorder. Multiple genes might interact with environmental factors to increase the risk of cataracts. Gene polymorphisms, including kinesin light chain 1 (KLC1), apolipoprotein E (APOE), xeroderma pigmentosum complementation group D (XPD), X-ray cross-complementing group 1 (XRCC1), glutathione S-transferase T1 (GSTT1), might also increase susceptibility to environmental influences (90–93).

1.4 Clinical Presentation and Diagnosis of Cataracts

A cataract is a chronic ocular disease. It develops slowly and may be asymptomatic at the early stages (94). With time cataracts could present with a variety of symptoms and signs. The most

common symptoms are blurred or clouded vision, night blindness, glare sensitivity, double vision and a frequent change in eyeglass prescription (95,96). Reduced visual acuity is a major sign of cataracts, which could influence reading at all distances (97).

Eye examinations may include visual acuity test, slit-lamp examination, slit-lamp photographs, and retinal exams. Among these examinations, slit-lamp examination and slit-lamp photographs are most commonly used in the clinical setting (1,2). To standardize the assessment of cataracts, several cataract grading schemes have been implemented. Lens Opacities Classification System III is widely used now, allowing ophthalmologists to diagnose cataracts at an early stage, and better assess, grade and monitor progression in a standardized way (98). Other grading systems, including the Oxford Clinical Grading system and Wisconsin cataract grading system, have also been used in previous studies (98).

1.5 Treatment of Cataracts

A cataract is an irreversible degeneration of the lens. Thus, the goal of cataracts treatment is to slow the progression and preserve vision to improve the quality of daily living. This section lists the main non-surgical treatment (section 1.5.1) and surgical treatment (section 1.5.2).

1.5.1 Non-surgical Treatment

Patient education is of importance and should start as early as patients are diagnosed with a cataract. Lifestyle modification, such as stopping alcohol, quitting smoking, regular exercise and more fruit and vegetable intake, may slow cataract progression. For patients with a visual acuity of 6/24 or above, eyeglasses or contact lenses can be used to correct refractive errors. Other options may include pupillary dilation (94).

1.5.2 Surgical Treatment

When visual functioning due to cataracts is greatly affected and everyday needs cannot be met, a patient is advised to proceed with cataract surgery. Cataract surgery includes two approaches: intracapsular extraction and extracapsular extraction. The intracapsular technique involves removing the entire lens, including the capsule. As intracapsular extraction has a poorer visual consequence and higher occurrence of complications, it is barely used in the developed world (1,99). Extracapsular extraction refers to removing the lens while the posterior and equatorial

capsule is retained within the eye. The remained capsule plays a role in containing replacement lens implantation (1,99). Phacoemulsification, one specific extracapsular extraction, is a modern cataract surgery that utilizes an ultrasonic device to break the cloudy lens down into small pieces that can be aspirated and replaced by an artificial lens. The advantages of phacoemulsification are a small incision and fewer complications (99). Thus, it results in quicker visual rehabilitation than other surgical procedures.

1.6 Prognosis of Cataracts

The prognosis of a cataract depends on a variety of factors, such as the age of intervention, intervention methods, types of cataracts, grading of lens opacity and co-existing other ocular diseases or systematic diseases. In general, thanks to advancements in modern technology, modern cataract surgery is safer and more effective than decades ago due to fewer complications and significant improvements in visual acuity. Roughly 85%-90% of eyes achieve best-corrected visual acuity of 6/12 (20/40 or 0.5) (99). Studies have shown that cataract surgery dramatically improves the quality of life and reduces all-cause mortality (100,101).

Chapter 2 Literature Review

2.1 Overview of Literature Review

The review of relevant literature is structured into two sections. Section 2.2 reviews the effect of alcohol on cataracts. Section 2.3 reviews the impact of fruit and vegetable (FV) consumption on cataracts.

2.2 The Impact of Alcohol on Cataracts

This section provides a review of the literature regarding the impact of alcohol intake on cataracts.

2.2.1 Search Strategy and Study Selection

I electronically searched PubMed, EMBASE, Web of Science and Cochrane Central Register of Controlled Trials comprehensively from January 1960 to December 2019. Search terms were as follows: (drinking OR alcohol OR ethanol OR lifestyle) AND (cataract OR lens opacity OR lens opacification). The search was limited to studies in humans. In addition, the references list of relevant studies and reviews were manually screened for eligible articles and searches were also conducted for studies that cited other relevant articles.

I included studies of any study design if they reported an association between alcohol use and age-related cataracts and/or cataract extraction. Studies with participants younger than 19 years old were excluded. The detailed search results are illustrated in the preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow chart (Figure 1) (102).

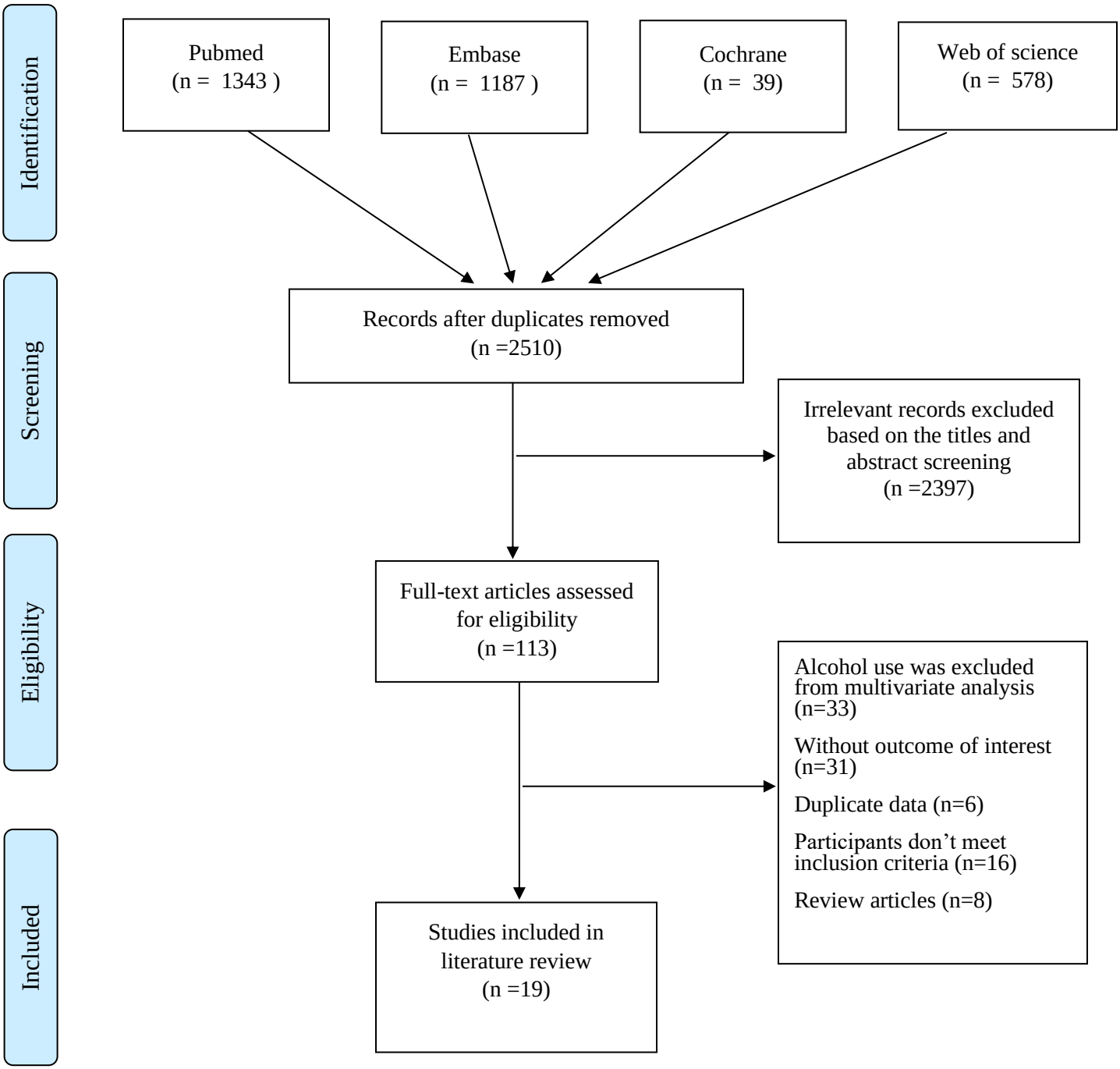


Figure 1. PRISMA flow chart of the review process for study objective 1

2.2.2 Findings from Literature Review Regarding Alcohol Intake and Cataracts

This section details the findings from the included studies regarding the association between alcohol consumption and cataracts. A list of study characteristics is provided in Appendix A.

2.2.2.1 Cohort Studies

A prospective population-based cohort study that used data from the Swedish Mammography Cohort (n=34,713, ages 49 to 83 years) examined the association between alcohol use and cataract extraction. All the participants were asked about the frequency and amount of alcoholic beverages they consumed. The total amount of alcohol was derived based on the answers and converted to grams per day. Alcohol use was divided into three groups: never drinkers, past drinkers, and current drinkers. Current drinkers were further classified as: <6, 6-13, >13-20, >20-30, >30 g/day. Cataract extraction was identified through record linkage. The study used the Cox proportional hazards model to estimate the hazard ratio while controlling for the following variables: age, smoking, diabetes, hypertension, use of steroid medication, use of antioxidant supplementation, BMI, and education. The analysis showed that current drinkers were more likely to have cataract extraction (HR:1.11, 95% CI: 1.02-1.21) than never drinkers, while this trend was not shown among past drinkers. Among current drinkers, the increase of 1 drink per day (13 g of alcohol use) was associated with a 7% increase in the hazard of cataract extraction (HR: 1.07, 95% CI: 1.02-1.12) (Authors used “rate ratio” in paper) (103).

Floud et al. conducted a prospective cohort study among postmenopausal women (average 56 years old) in the United Kingdom, using data from the Million Women Study. A total of 1,312,051 women were followed for a mean period of 10.7 years. Alcohol intake and other covariates were ascertained at recruitment through a structured interview. Cataract surgery, emigration and death were assessed by linked central registers. A Cox proportional hazards regression model was fitted to estimate the hazard ratio adjusted for age, residence, education, smoking, BMI, physical activity, diabetes, age at menarche, parity, oral contraceptive use and hormone therapy. Alcohol use of 2-14 units per week (one unit=10g alcohol) had a lower risk of cataract surgery (HR: 0.93; 95% CI: 0.91-0.94), while heavy drink (more than 15 units per week)

conferred a HR of 0.92 (95% CI: 0.89-0.96), in comparison to the consumption of fewer than two units per week (Authors used “relative risk” in paper) (4).

The Salisbury Eye Evaluation study was a population-based longitudinal study in Salisbury, Maryland. It enrolled 2,520 people aged 65 to 84 years old and followed them for up to two years. This study aimed to examine racial differences in lens opacity incidence and progression. Alcohol consumption was assessed through an administered questionnaire at baseline and participants were classified as never drinkers, past drinkers, and current drinkers. Lens photographs were taken to examine cataracts. At two years of follow-up, 625 participants were excluded from the analysis due to bilateral cataract surgery at baseline, loss to follow-up, or no available lens photographs. There was a statistically significant association between participants and non-participants in terms of age ($P<0.0001$), smoking ($P=0.04$) and steroid use ($P=0.0001$). After adjusting for potential confounders, multiple logistic regression models showed that there was no statistically significant association between past or current alcohol drinking and cortical cataracts, with an odds ratio (OR) of 1.45 (95% CI: 0.74-2.86) and 1.33 (95% CI: 0.70-2.49, respectively. The odds of nuclear cataracts was also not statistically significantly associated with past (OR: 0.83, 95% CI: 0.50-1.37) or current drinking (OR: 0.90, 95% CI: 0.59-1.37) (5).

Kuang et al. conducted a cohort study using the Shihpai Eye Study cohort to report the 7-year incidence of cataracts among people aged 65 or older in Shihpai, Taipei. Of 2,405 participants randomly selected from 4,750 residents of the baseline survey in 1999, 1361 participated in both the questionnaire and eye examination. Of the 1,361 participants at baseline, 460 subjects were eligible for follow up. Non-participants were more likely to be older, female, without a spouse, lower education and never smokers. A structured questionnaire was conducted to obtain information regarding alcohol consumption. Alcohol intake was only limited to wine and hard alcohol, and alcohol consumption was categorized into two types: yes and no. The frequency of alcohol consumption of more than once a week was coded as alcohol drinking, while the frequency of alcohol consumption of only once or less per week was considered no consumption. Cataracts were ascertained through an eye examination by ophthalmologists. Generalized estimating equations were used to fit the models. Multivariable analysis showed a negative association between alcohol intake and cortical cataracts, although it was not statistically significant (relative risk [RR]:0.53, 95% CI: 0.26-1.07) (104).

The Blue Eye Study by Kanthan et al. was a population-based cohort study, including 3,654 participants aged 49 years and older at recruitment in the Blue Mountains region, Australia. Two follow-ups were carried out in five years and ten years, with 2,335 and 1,952 subjects returning for examinations, respectively. An interviewer-administered questionnaire was used to assess alcohol consumption. The alcohol intake was assessed for frequency (days per week), amount (drinks per day) and types (beer, wine or spirits). Alcohol consumption was categorized as 0, >0 to ≤1, >1 to ≤ 2, and >2 drinks per day. Lens photographs were performed to examine types of cataracts. An initial discrete logistic model adjusted for age and gender was fitted to estimate the association of alcohol consumption with incident cataracts. Potential confounders, including diabetes, steroid use, myopia, socioeconomic status, and smoking, were included for further adjustment if they showed statistically significant associations with cataracts. Moderate alcohol consumption (>1 to ≤ 2 drinks per day) was used as the logistic regression model's reference level. White wine was found to have a significant association with incident posterior subcapsular cataracts (PSC). The multivariable-adjusted odds ratio was 2.75 (95% CI: 1.01-7.41) for light alcohol drinking (>0 to ≤1 drink per day). However, the association was not statistically significant between abstainer or heavy drinker and PSC. Interestingly, no statistically significant association was found between consumption of beer, spirits, red wine and white wine and incidence of nuclear and cortical cataracts. It was also reported that aggregated alcohol consumption was not statistically significantly associated with the three cataract subtypes (105).

The Beaver Dam study was a population-based longitudinal study conducted among persons aged 43 to 86 years old in Wisconsin. The primary aim of this study was to examine the association of socioeconomic and lifestyle factors with cataracts. A total of 4,926 participants were enrolled from 1988 to 1990 and followed from five years to ten years. The final analysis included those who completed all three visits (n=2,764) or those who completed the baseline examination and a 5-year follow-up (n=920). The non-participants and participants had different baseline distributions of age, education, income, smoking, hypertension, visual acuity, blood pressure and cholesterol. The information regarding alcohol consumption was collected through a survey, and the response was converted to grams of ethanol based on the servings of beer, wine or liquor consumed in the past year. Graders were masked to examine cataracts via lens photographs. The amount of alcohol intake was classified into four categories (0-39g, 40-99g, 100-334g, ≥335g) in the analysis. The study showed that consumption of more than 335g ethanol

was statistically significantly associated with the incident nuclear cataracts (HR: 1.93; 95% CI: 1.08-3.46) compared with never drinking, while the association between alcohol use and cortical cataracts was not significantly associated. Interestingly, consumption of 0-39g and 40-99g alcohol had an HR of 0.52 (95% CI: 0.31-0.86) and 0.46 (95% CI: 0.27-0.80) (42).

A study by Manson et al. using data from the prospective data of Physician's Health Study explored the impact of alcohol consumption and risk of cataracts and cataract extraction. This study included 17,824 male physicians with 88,565 person-years of follow-up. The estimated relative risk (RR) of incident cataracts was 1.31 (95% CI: 0.95-1.81) for subjects with daily alcohol consumption, compared with those who only used alcohol less than once per month (106).

Chasan-Taber et al. conducted a cohort study using the Nurses' Health Study data. This study was carried out between 1980 and 1992, and information was collected via biennial questionnaires. The aim of this study was to examine the association between alcohol intake and the incidence of cataract extraction. A total of 77,466 women were included with 761,036 person-years of follow-up. Alcohol use was computed as average daily intake (grams/day) based on the food frequency questionnaire. To reflect the impact of long-term alcohol intake, the study updated daily alcohol use based on the cumulative consumption prior to the start of each follow-up interval. Cataract extraction was collected via questionnaire and ascertained by patients' ophthalmologists. Alcohol consumption was categorized as: almost never use, ≤4.9, 5.0-14.9, 15.0-24.9 and ≥25 (g/day) in the analysis. A Cox proportional hazards model adjusted for potential confounders was fitted to estimate the effect. The following variables were adjusted for: time period, physical activity, smoking, parental history of myocardial infarction, area of residence, BMI., number of physician visits, aspirin use, energy intake, cholesterol level, diabetes and hypertension. . The hazard ratio (HR) was 0.84 (95% CI: 0.73-0.96) for the alcohol use of ≤4.9 (g/day), 0.83 (95% CI: 0.71-0.97) for the alcohol use of 5.0-14.9 (g/day), 1.08 (95% CI: 0.87-1.33) for the alcohol use of 15.0-24.9 (g/day), and 1.10 (95% CI: 0.74-1.62) for the alcohol use of ≥25 (g/day), respectively (Authors used “relative risk” in paper) (107).

A recent cohort study was conducted among adults aged 50 years or older in Kenya using data from the Nakuru Eye Disease Cohort Study. The study aimed to estimate the six-year cumulative incidence of cataracts and examine risk factors. The alcohol use was divided into never drinking,

past drinking and current drinking. After adjusting for potential confounders, a Poisson regression analysis showed that current alcohol drinkers had a higher relative risk (relative risk [RR]:1.4, 95% CI:1.1-1.8), compared with never drinkers, while the RR for the former drinkers was 1.1(95% CI: 0.9-1.3) (6)

2.2.2.2 Cross-sectional Studies

Nam et al. conducted a cross-sectional study using data from the 2008-2011 Korea National Health and Nutrition Examination Survey (KNHANES). A total of 16,014 subjects aged ≥40 years were enrolled based on a stratified, multistage and cluster sampling method. After removing 148 subjects due to missing value, 15,866 participants were included in the final analysis. All lifestyle factors and sociodemographic characteristics were collected through an interview or self-report questionnaires. Alcohol intake was classified into three groups: non-drinker, light to moderate drinker (1-30g/day) and heavy drinker (>30g/day). A comprehensive eye examination was conducted by ophthalmologists to assess lens opacities. Two separate logistic regression models for men and women were fitted to examine the association of various covariates with cataracts. The following variables were adjusted in the models: age, household income, education, alcohol use, smoking, physical activity, BMI, diabetes, hypertension, history of cardiovascular disease or stroke, hormone replacement therapy (women), occupation, residential area, and sun exposure. The results showed that the association of alcohol consumption with any cataracts was statistically insignificant. The odds ratio (OR) was 0.92 (95% CI: 0.75-1.14) for moderate drinkers, and 1.12 (95% CI: 0.87-1.44) for heavy drinkers in men. The OR was 1.06 (95% CI: 0.92-1.23) and 1.00 (95% CI: 0.56-1.80), respectively in women(108).

Li et al. conducted a cross-sectional study among 8,445 Chinese aged 18-94. Alcohol consumption was collected via a structured face-to-face interview. Participants were categorized as never drinkers, previous drinkers, and current drinkers. Participants who had at least one drink of alcoholic beverage during the past 30 days of the interview date were considered current drinkers. Information regarding alcoholic beverage consumption was collected for the past 30 days, and alcohol use was derived from the collected data. The amount of alcohol intake for current drinkers was categorized as low intake (≤1 drink/week), moderate intake (1-14 drinks/week) and heavy intake (>14 drinks/week). Ocular examinations were conducted by

groups of ophthalmologists who received special training and followed a standardized procedure. A multivariable logistic regression model adjusted for age and sex was used to estimate an odds ratio. Compared with abstainers, the OR (95% CI) for former drinking was 1.0 (0.3-2.6). For current drinkers, the amount of drink was not significantly associated with cataract. The OR for ≤1 drink/week, 1-14 drinks/week, and >14 drinks/week was 0.8 (95% CI:0.1-4.8), 0.9 (95% CI:0.4-2.4) and 1.2 (95% CI:0.4-3.6), respectively (109).

The Singapore Malay Eye Study was a population-based cross-sectional study, which included 3280 participants(ages 40 to 80 years) out of 4168 eligible subjects (78.7% response rate). Alcohol use was categorized into drinkers and non-drinkers. Lens opacities were assessed through a slit lamp by an ophthalmologist. The adjusted logistic regression model was used to estimate the association between potential risk factors and cataracts. The study indicated that alcohol intake was not statistically significantly associated with cataract prevalence (7).

Morris et al. examined the association of alcohol intake with lens opacity using the Nurses' Health Study cohort in a sample of 556 women aged 53-74 years. Alcohol consumption was obtained via food frequency questionnaires. Both the frequency and amount of alcohol use for three types (spirits, wine and beer) were recorded. Average daily alcohol intake was included as a continuous variable in the final analysis. A detailed eye examination was undertaken to assess the lens opacity. The degree of lens opacity was determined by the Lens Opacities Classification System III(LOCS III). The 60^{th} percentiles of the distributions of nuclear and cortical opacity LOCS III grades were used as the cut-off point of higher vs lower opacity. Multiple logistic regression analysis showed a 10g increase of average daily alcohol consumption was associated with the increased odds of nuclear opacity (OR: 1.30; 95% CI: 1.10-1.54) after controlling for age and vitamin C. In contrast, the association was not significant for cortical opacity (OR: 0.9; 95% CI: 0.8-1.0) (110).

Tsai et al. conducted a population-based cross-sectional study among 1361 residents aged 65 years or older in Shi-Pai, Taiwan region. This study aimed to examine the prevalence and risk factors of age-related cataract. Participants with congenital and traumatic cataracts were excluded from the study. A structured questionnaire was used to collect information regarding demography, lifestyle factors and past medical history. Alcohol intake was grouped as drinkers (more than once per week) and non-drinkers (once per week). Cataracts were assessed by senior

ophthalmologists using the LOCS III system. The following variables were included as independent variables: waist-to-hip ratio, age, gender, blood pressure, history of Diabetes, hormone use for women only, alcohol use and smoking. Logistic regression analysis showed that alcohol intake was not significantly associated with the prevalence of three types of cataracts: nuclear (odds ratio [OR]: 1.0; 95% CI: 0.7-1.3), cortical (OR: 0.8; 95% CI: 0.6-1.2), and posterior subcapsular cataracts (OR: 0.8; 95% CI: 0.5-1.2) (111).

McCarty et al. conducted a cross-sectional study, including 5141 participants aged 40 and older in Australia. The study aimed to examine the prevalence and risk factors of cataracts in Australia. Lens opacities were examined and graded during the scheduled examination. Information regarding potential risk factors was collected via a standardized questionnaire. Alcohol use was categorized as current drinkers, past drinkers and never drinkers based on the drinking status and was divided into none, ≤1, ≤2 and >2 drinks /day based on the amount of consumption. Pearson chi-square test was used to assess all categorical variables. The study showed a significant association between alcohol intake for those with > 4 drinks/day and cortical cataract (p=0.001). In comparison, an intake of >2 drinks/day was significantly associated with the prevalence of nuclear cataract (p=0.001). The association was found to be significant between alcohol use and posterior subcapsular cataract among past drinkers as well (p=0.016) (112).

Cumming et al. examined the association of alcohol consumption, smoking with cataracts using data from the Blue Eye Study in Australia. A total of 3654 participants aged 49- 97 years old were included in the analysis. Cataracts were assessed by ophthalmologists, and the possible risk factors were collected via an interviewer-administered questionnaire. Information regarding alcohol use was collected through questions about the frequency (days/week), the usual number of drinks/day and type of consumed alcohol. Alcohol consumption was divided 4 categories(none, <1, 1-3 and ≥4 drinks/day). Ordinal regression was used to examine the association of alcohol with cortical and posterior subcapsular cataract. A logistic regression model was fitted to investigate the association between nuclear cataract and alcohol use. The study showed that the association between nuclear cataract and heavy alcohol use (≥4 drinks/day) was significant (odds ratio [OR]:2.1, 95% CI:1.1-4.3). In comparison, the

association between any alcohol intake with cortical and posterior subcapsular cataract was not significant (113).

A recent population-based cross-sectional study in India included 6,617 subjects (≥60 years) in urban and rural areas. The aim of this study was to examine the prevalence of cataracts and their associated risk factors. A total of 4,331 subjects (65.5%) were included in the final analysis after applying the exclusion criteria. A detailed questionnaire was used to collect information regarding alcohol consumption. Lens opacities were assessed by experienced ophthalmologists. Multivariable logistic regression analysis showed alcohol intake was not associated with cataracts in both rural (odds ratio [OR]:0.77, 95% CI:0.44-1.37) and urban populations (OR: 1.03; 95% CI: 0.48-2.21) (114).

2.2.2.3 Case-control Studies

A case-control study by Tavani et al. aimed to examine the relationship between food and nutrient intake and cataract risk was conducted in Italy, including 207 cases and 706 controls (ages 21 to 80 years) (115). Total alcohol intake was computed based on the collected information regarding alcoholic beverages consumption (wine, beers and spirits). Average daily alcohol consumption was categorized as non-drinkers, <3, 3-5, 5-8, and ⩾8 drinks/day in the analysis. Cataracts were diagnosed during hospitalization. A multivariable logistic regression model was fitted to estimate the effect of alcohol on cataracts, with adjusting for age, sex, education, smoking, diabetes, BMI and calorie intake The study showed that consumption of fewer than three drinks per day was associated with the risk of cataracts (OR: 1.6’ 95% CI: 1.0-2.4). In contrast, no statistically significant association was found for the other alcohol intake groups (115).

Another matched case-control study conducted by Phillips et al., which enrolled 990 cases and 858 controls, and indicated that ‘light and frequent’ and ‘light and infrequent’ alcohol use was statistically significantly associated with the lower risk of cataracts, compared with non-drinkers and 'occasional' drinkers, showing a U-shaped association (116).

2.2.3 Summary of Included Studies

Through our systematic literature review, a total of 19 studies were identified as meeting our inclusion criteria. Two of them were case-control studies; eight of them were cross-sectional studies and nine cohort studies. However, the findings across the studies on the association between alcohol use and cataract were inconsistent. Furthermore, many studies have significant limitations, which could distort the validity of estimation.

Among eight cross-sectional studies were included in our search, three studies found a statistically significant association of alcohol use with cataracts. The studies by Morris et al. and by Cumming et al. reported a positive association of alcohol use with the risk of nuclear opacity, but not for other cataract types (110,113). The study by McCarty et al. found a statistically significant association for all three cataract types (112).

The alcohol measurement varied substantially across the studies. Morris et al. treated alcohol use as a continuous variable (110). The remaining studies treated alcohol use as a categorical variable but used different cut-points for categories. For example, Wu et al. simply grouped participants into drinkers and non-drinkers (7). Nam et al. divided participants into non-drinker, light to moderate drinker (1-30g/day) and heavy drinker (>30g/day). The difference in alcohol measurement might partially explain the inconsistent findings (108).

It should be noted that the temporality of association raises a concern when we try to make causal inferences using cross-sectional data since alcohol use and cataracts were measured at the same time, meaning alcohol use could precede cataracts, or cataracts could precede drinking habits. Thus, we cannot infer that alcohol causes cataracts given the evidence from cross-sectional studies.

Of nine cohort studies, four studies reported a positive association between alcohol intake and cataracts, while no statistically significant association was found among the other four studies (4,5,42,51,105–107,117). Interestingly, one study by Floud et al. indicated that alcohol consumption was negatively associated with cataract surgery (4). The difference in effect estimation could arise due to the presence of different degrees of biases in these studies.

Loss to follow-up is not uncommon in the cohort studies, which could introduce selection bias, especially when the percentage of loss to follow up is considerable. For example, the studies by Klein et al. and Bastawrous et al. had a follow-up rate of 56% and 53%, respectively (42,117). The authors addressed attrition in an inappropriate way (e.g. exclude records with missing data), and considerable attrition among these studies could seriously compromise the internal validity of the estimates.

All of the included cohort studies categorized alcohol intake based on drinking status, frequency or average amount of alcohol use except for a study by Lindblad et al. (103). For instance, one study compared daily alcohol use with less than once per month (106). Alcohol intake in another study was only classified into two categories: yes and no (51). Three studies divided drinkers into never drinker, former drinker and current drinker (5,103,117). Furthermore, non-drinker was defined differently across studies. For example, Manson et al. treated less than once per month as non-drinkers, and Kuang et al. considered less than once per week as non-drinkers (51,106). In general, categorization may result in loss of power, precision, and model misspecification; in addition, it complicates synthesizing research evidence across studies. Categorization is advised to be avoided when continuous data are available (118).

One case-control study reported a U-shaped association (116), while another study showed a statistically non-significant association (115). Similarly, both of them categorized alcohol intake, which may have produced imprecise estimation. Additionally, subjects with cataracts may have been more likely to recall alcohol use, thus recall bias could be introduced.

None of these studies dealt with time-varying covariates in the analysis. All the risk factors were measured at baseline and included in models. However, drinking status and some other covariates (e.g. BMI and history of chronic disease etc.) may change over time during the follow-up period, and ignoring the time-varying nature of these covariates could distort the estimation of the effect of alcohol use on cataracts. Most studies addressed missing data in an inappropriate way. For example, they excluded subjects with incomplete information. If subjects with incomplete information constituted a large percentage of the study population, which is common, simple elimination could incur biased estimates and reduce statistical power.

2.3 The Association of FV Consumption with Cataracts

This section provides a comprehensive literature review regarding the association between FV consumption and cataracts.

2.3.1 Search Strategy and Study Selection

This systematic review was performed using PubMed, EMBASE, Web of Science and Cochrane Central Register of Controlled Trials from January 1960 to February 2020. Keywords were "lens opacity", "lens opacification", "cataract", "fruit", "vegetable", "diet". The search was restricted to the English language and to studies in humans.

A two-level screening was performed. Titles and abstracts were used to perform the first level screening. Any studies that included diet or any specific food items as exposure variables and cataracts or cataract surgery as outcome variable was retained. Duplicates and irrelevant studies were removed, which resulted in 31 full-text papers for the next level screening. For the second level screening, only the studies examined the association of fruit and/or a vegetable intake with cataracts/cataract surgery were included. We excluded studies including subjects younger than 19 years old. We also excluded any reviews, editor letters, commentaries, case reports and conference abstracts. The detailed search results are illustrated in the PRISMA flow chart (Figure 2) (102).

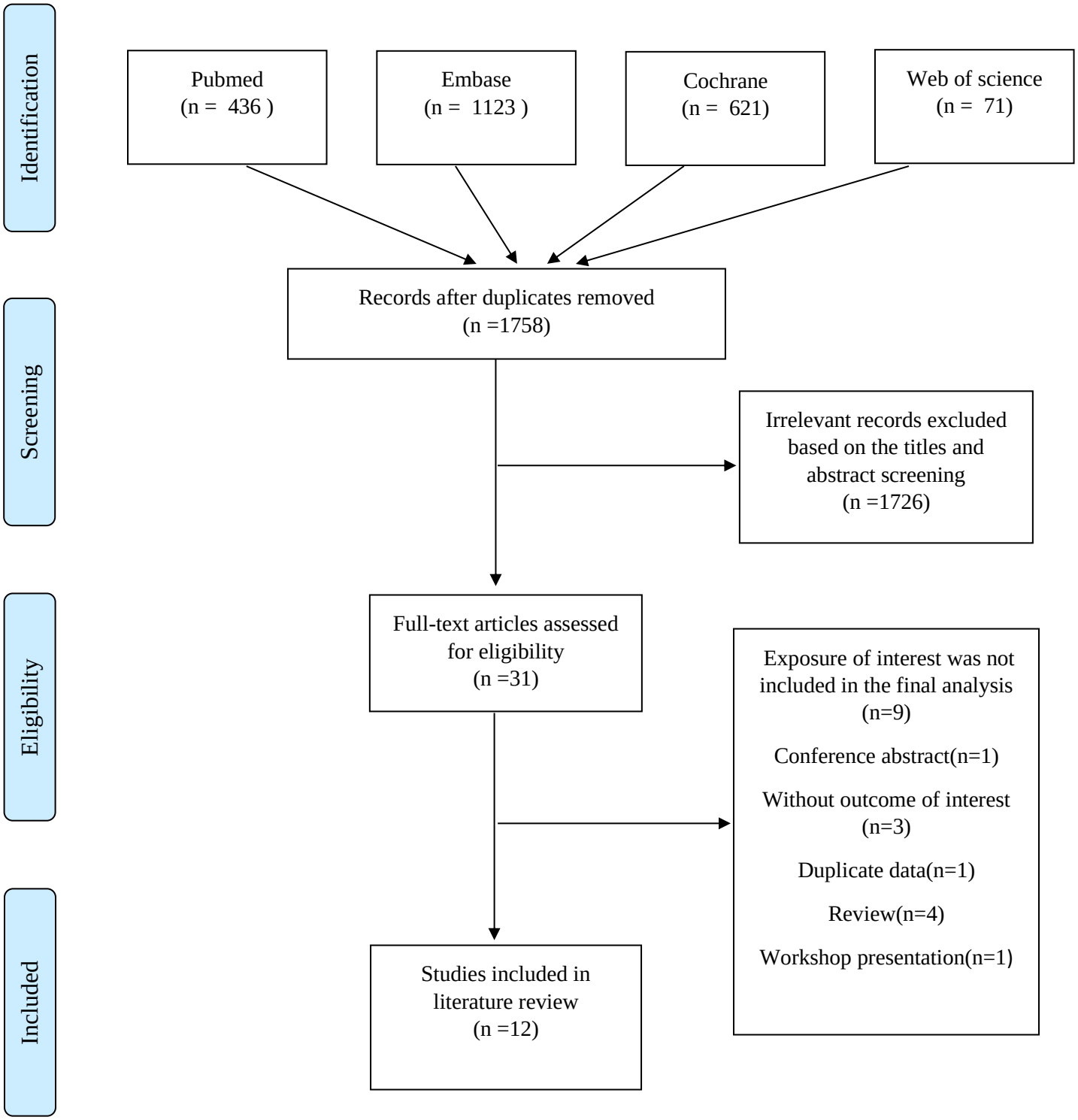


Figure 2. PRISMA flow chart of the review process for study objective 2

2.3.2 Findings from Literature Review Regarding Fruit and Vegetable Consumption and Cataracts

Fruit and vegetable consumption have been suggested as a protective factor in several observational studies. Oxidative stress has been proved to play an essential role in cataractogenesis (119). Thus, it is reasonable to hypothesize that antioxidants in most fruit and vegetables may prevent the development of cataracts. This section presents literature based on the study design. A list of included studies is provided in Appendix B.

2.3.2.1 Cohort Studies

Adachi et al. conducted a population-based cohort study using data from the Japan Public Health Center-based Prospective (JPHC) Study among 71,720 subjects aged 45-74 years old in Japan. Those with missing information on FV intake and a history of cancer, diabetes and cardiovascular disease were excluded from the analysis. Separate questionnaires were administered at baseline, 5-year follow up and 10-year follow-up. This study used a food frequency questionnaire (FFQ) to collect FV consumption, including 30 vegetable and 16 fruit items. Frequency and portion size were collected for each food item, and daily consumption was computed by multiplying frequency and portion size. A multiple logistic regression model was fitted to estimate the association of FV with incident cataracts. The following confounders were adjusted in the analyses: BMI, smoking, alcohol intake, vitamin intake, and history of fundus photographic examination. This study found that total vegetable consumption was not statistically significantly associated with cataract in males (OR: 0.77; 95% CI: 0.59-1.01) but was statistically significantly associated with cataracts in females (OR: 1.28; 95% CI: 1.06-1.53). However, fruit intake was not statistically significantly associated with cataracts for both males and females (120)

A population-based cohort study using the Swedish Mammography Cohort (SMC) examined total antioxidant capacity of the diet and cataracts among women aged 49-83 years old with a mean of 7.7 years follow-up. In this study, subjects with a history of heart disease, stroke and diabetes were excluded from SMC, which left a total of 36,707 women at the start of follow-up.

Dietary total antioxidant capacity was obtained by multiplying the mean frequency of each food item intake by oxygen radical absorbance capacity content. Fruit and vegetables contributed a significant part of the total antioxidant capacity of the diet (44.3%). Total antioxidant capacity of the diet was categorized based on quartiles. Cataracts and cataract surgery were ascertained and cross-validated through various registers. The hazard ratios (HR) were used to estimate effects by fitting a Cox proportional hazards regression model adjusted for age, smoking, waist circumference, educational level, hypertension, hormone therapy, corticosteroid use, dietary supplement use, alcohol consumption, and energy intake. The HR was 0.87 (95% CI: 0.79-0.96) when comparing the highest quintile of the total antioxidant capacity of the diet with the lowest quintile. In further subgroup analysis, the effect estimate was stronger among women 65 years or younger, with a HR of 0.78 (95% CI: 0.65-0.94), while for the group older than 65 years, the HR was 0.92 (95% CI: 0.82-1.04). The p-value for interaction between total antioxidant capacity and age was 0.07. The association between total antioxidant capacity and cataracts did not significantly differ across categories of smoking and waist circumference (P-value for interaction was 0.39 and 0.52, respectively) (121).

Another prospective cohort study performed by Christen et al. included 35,724 female health professionals aged 45 years or older. The aim of this study was to examine if higher consumption of fruit and vegetable reduce the incidence of cataracts and cataract extraction. Total fruit and vegetable intake were calculated by summing the average daily intake of each food item acquired from a validated, semiquantitative food-frequency questionnaire that included 29 vegetables and 15 fruit items. The fruit and vegetable intake were categorized based on the quintiles. The 1^{st} quintile and 5^{th} quintile were 2.6 servings/day and 10 servings/day, respectively. Cataracts or cataract extraction was collected through questionnaires and confirmed by medical records subsequently. Hazard ratios were estimated by fitting a Cox proportional hazards regression model adjusted for age, randomized treatment assignment, smoking, alcohol use, history of diabetes, history of hypertension, history of hypercholesterolemia, BMI, physical activity, parental history of myocardial infarction, menopausal status, postmenopausal hormone use, use of multivitamins or vitamin C supplements, total energy intake, and history of an eye exam in the past two years. The aggregate fruit and vegetables intake were shown to slightly reduce the risks of cataracts (hazard ratio [HR] for 5^{th} quintile vs 1^{st} quintile =0.83, CI 95%: 0.79-0.99). However, the analysis conducted for fruit and vegetable separately indicated that the association

was not statistically significant. Further investigation for interaction between fruit and vegetable and cataracts was not statistically significant as well (Authors used “relative risk” in paper) (122).

Chasan-Taber et al. conducted a prospective cohort study among 50,461 female nurses aged 45-71 years old, including cohort members of the Nurse's Health Study cohort in the United States. The aim of this study was to investigate the association between carotenoid and vitamin A intake and the risk of cataract extraction. The dietary intake was collected from 1980 to 1992 through a semiquantitative questionnaire biennially. Women who reported cancer and were younger than 45 years in 1980 were excluded from the cohort. Respondents were asked to report if they had a cataract extraction every two years from 1984 to 1992, and the responses were then confirmed by ophthalmologists. Cases that were considered as congenital or secondary cataracts by physicians were excluded. The assessment of individual foods rich in carotenoid was only carried out on 1980, 1982 and 1984 questionnaires. Multivariate models were fitted by adjusting for age, time period, diagnosis of diabetes, cigarette smoking, BMI, area of residence, number of physician visits, aspirin use, total energy intake, and alcohol intake. The study reported that spinach and other greens rich in carotenoids were associated with a decreased cataract extraction. Intake of 2 times/week of spinach and other greens reduced cataract extraction by 18% compared with consumption of less than one time/month (relative risk [RR]: 0.82, 95% CI: 0.68-0.98). Cooked spinach had an RR of 0.62 (95% CI: 0.45-0.86). However, other fruit and vegetables, such as apples, oranges, alfalfa, and cauliflower, were not statistically significantly associated with cataract extraction (123).

2.3.2.2 Case-control Studies

A case-control study was conducted in Greece, which included 314 cases and 314 frequency-matched controls aged from 45-85 years old. The case was defined as any cataracts diagnosed by an ophthalmologist. Controls were identified among visitors of patients visiting the ophthalmologic clinic. All cases and controls were asked to report the frequency of each of 120 food items included in a semi-quantitative food-frequency questionnaire. Potential confounders were adjusted, including age, sex, education, BMI, and smoking. This study showed that the odds of having cataracts were decreased by 53% (OR: 0.47, 95% CI: 0.38 to 0.59) and 47% (OR:

0.53, 95% CI:0.39 to 0.72) respectively for those consumed vegetables per 56.19 times/month and fruits per 109.41 times/month (124).

Ghanavati et al. conducted a case-control study in Iran, including 97 cases and 198 controls. This study examined the association between dietary intake and cataracts. Information regarding food consumption was collected via a validated food frequency questionnaire (FFQ). A total of 147 food items included in this questionnaire were categorized into five food groups: vegetable, fruits, grains, milk, and meat. Each food group were assigned a score based on the healthy eating index with the Likert scale from 0 to 10. This study showed that the healthy-eating-index scores of vegetable and fruit were significantly higher among healthy individuals than patients with cataracts (10 vs 7.8 and 9.8 vs 7.1, respectively) (125).

Lu et al. performed a case-control study among 360 cases and 360 controls aged 45-85 years old. This study aimed to examine if a higher intake of fruit and vegetable was associated with age-related cataracts. Patients were those admitted to hospitals for lens implantation, and controls were from the same hospital with diseases unrelated to cataracts. After controlling for potential confounders, the study showed that intake of fruit, vegetables and a combination of fruit and vegetables reduced the risk of cataracts. In contrast to the lowest quartile of fruit intake in controls, the highest quartile of intake in cases increased the odds of cataracts by 19% (OR: 0.81, 95% CI: 0.67-0.97). The odds ratio was 0.81 (95% CI: 0.69-0.94) for vegetable intake and 0.71 (95% CI: 0.60-0.93) for total fruit and vegetable intake (126).

Another case-control study among 31 cases and 31 controls was conducted in Nigeria. The consumption of a wide range of food items, including fruit, vegetables, and animal and dairy products, were collected via a food frequency questionnaire. Cases and controls reported different frequency of fruit and vegetable intake per week. A Chi-square test was performed to compare the difference between cases and controls. While 39% of controls consumed fruits more than five times/week, this percentage was only 6.4% among patients with cataracts. There was a similar trend regarding the consumption of vegetables. The subjects consuming vegetables at least three times a week constituted 29% of controls, compared with only 9.6% in the case group. The difference in fruit and vegetable intake across the two groups was statistically significant (127).

Tavani et al. conducted a case-control study in Italy which included 207 cases and 706 controls aged 25 to 80 years old from the same health network hospitals. A total of 34 food items were collected in terms of frequency of consumption per week. Cases were those admitted to a hospital for cataract extraction. Controlled covariates were age, sex, education, smoking, diabetes, BMI and calorie intake. This study indicated that specific food items (cruciferae, spinach, tomatoes, and melon) had an inverse association with the cataract extraction. The odds ratios of the 3rd tertile group compared with the the 1st tertile group for cruciferae, spinach, tomatoes, and melon were 0.5 (95% CI: 0.3 to 0.8), 0.6 (95% CI: 0.4 to 0.9), 0.5 (95% CI: 0.4 to 0.8), and 0.5 (95% CI: 0.4 to 0.8) respectively, while the association for the remaining fruits and vegetable was not statistically significant (115).

2.3.2.3 Cross-sectional Studies

Pastor-Valero et al. conducted a cross-sectional study among 593 subjects aged 65 years and older from the European Eye Study. A validated semi-quantitative Food Frequency Questionnaire was used to assess dietary intake. Cataracts were diagnosed by an ophthalmologist. A logistics regression model was fitted to examine the association between FV consumption and cataracts. It indicated that aggregate fruit and vegetable intake was statistically significantly associated with the prevalence of cataracts or cataract extraction after adjusted by sex, age, marital status, smoking, alcohol drink, physical activity, supplement use, energy intake, obesity and diabetes (P-trend=0.008). The highest quartile and the third quartile of fruit and vegetable intake reduced the cataracts risk by 62% (OR=0.38, 95% CI: 0.20 to 0.70) and 55% (OR=0.45, 95% CI: 0.24 to 0.84) respectively compared with the lowest quartile (128).

Another cross-sectional study among 500 black diabetic patients in Congo was conducted to examine if fruit and vegetable intake reduced the risk of cataracts. Fruit and vegetables rich in antioxidants (e.g. Brassica rapa, Phaseolus vulgaris, Abelmoschus spp, and Musa acuminata) were included in this study. Regular intake of these food items was compared with never intake. The Chi-square test indicated that regular intake of each food item was statistically significantly associated with cataracts among type 2 diabetic patients (129).

The Blue Mountains Eye Study was a population-based cross-sectional study of eye diseases among 2900 participants aged 49 to 97 years in Australia. This study examined the association of

various nutrients and vegetables with three types of cataracts. A semi-quantitative questionnaire was used to acquire information on dietary intake. The vegetable consumption was categorized into five categories (<1 time/month, 1-3 times/month, 1 time/week, 2-4 times/week, >4 times/week) based on the intake frequency. Through age and sex adjusted logistic regression models, the study showed that the association between various vegetables and nuclear cataracts were not statistically significant (130).

2.3.3 Summary of Included Studies

Of 12 identified studies from our systematic literature review, three studies were cross-sectional studies. One of the concerns for the cross-sectional studies was that information regarding exposures and outcome were collected at the same time, which limited the inference of temporal association between risk factors and outcome. For example, dietary habits may impact the development of cataracts, while participants who suffered from cataracts may change their nutritional habits. In addition, the study conducted by Pastor-Valero et al. had a low response rate (50%). No information was obtained from the non-respondents, and a bias may have been introduced (128). The study by Moise et al. did not adjust for confounders, and the estimation may have been distorted by potential confounders (129). The (non-statistically significant) association between exposure and outcome in the Blue Mountains Eye Study may be partially attributed to unmeasured confounder variables since only age and sex were adjusted in their analysis (130).

The case-control study conducted by Theodoropoulou et al. did not control for some important confounders, namely, diabetes, income and alcohol drinking (124). Uncontrolled confounders may distort the estimation and bias causal inference. Tavani et al. used hospital controls from which cases were identified, which means controls could have suffered from diseases related to the exposure being studied, resulting in selection bias (115). Due to the nature of the case-control study design, differential misclassification bias could have been introduced when the fruit and vegetable consumption was collected through a food frequency questionnaire after the cataracts status was determined.

Four cohort studies were identified in our search. However, each study had some limitations. The two cohort studies conducted by Rautiainen et al. and Christen et al. categorized the exposure

variable based on quintiles, which might lead to inaccurate estimation because the assumption of homogeneous risk within each category is unrealistic (121,122). Moreover, the estimated results were non-comparable across studies because of various data-driven cut points used for categorization (131). Another limitation was that the exposure of interest and covariates were only measured once at baseline. However, some covariates (i.e., dietary intake, BMI etc.) may change substantially over time, which would distort the observed association. In the cohort study conducted by Chasan-Taber et al., cataract extraction was used as a proxy measure of outcome, which may underestimate the cases and lead to a biased estimate of exposure-disease association (123). Misclassification bias could also occur in our cohort studies. However, the misclassification of exposure status was likely non-differential. Thus, this type of bias would attenuate the estimates. Christen et al. only adjusted age and treatment, and unadjusted-for confounders could threaten the validity or effect estimation (122). None of the studies incorporated time-varying covariates. Furthermore, most studies excluded participants with missing data, apparently without checking the missing mechanism. More cohort studies are needed to fully adjust covariates and incorporate time-varying covariates and address the missing data in an appropriate way.

.

Chapter 3 Study Rationale and Objectives

This chapter presents the study rationale (section 3.1) and overview of the study objectives and hypothesis (section 3.2).

3.1 Study Rationale

The incidence of cataracts has grown significantly and become the leading cause of blindness globally. Therefore, identification of modifiable risk factors is of critical importance for prevention and early intervention Our current study aimed to investigate the association of alcohol use and fruit and vegetable consumption with cataracts. The previous chapter provided a comprehensive literature review regarding the association of alcohol use and fruit and vegetable intake with the risk of cataracts. However, the findings among these studies were inconsistent. Furthermore, many studies had significant limitations, which have been summarized in the previous chapters. For instance, most studies categorized alcohol use or fruit and vegetable intake into two or more groups, and none of the studies evaluated the non-linear association of these exposures with cataracts. There are no studies that investigated a cumulative and time-varying effect of alcohol or fruit and vegetables on cataracts. Many studies addressed the missing value in a potentially inappropriate way. None of the studies considered the uncertainty of estimation caused by missing data, even though the missing data comprised a large proportion of their sample population and missing did not follow a random manner. It is essential to better understand how alcohol use and fruit and vegetable intake affect cataracts so that effective health policies can be implemented to reduce the risk of cataracts. To address this knowledge gap, we used the longitudinal NHPS data to investigate the relationship between alcohol use and fruit and vegetable consumption and cataracts.

3.2 Study Objectives and Hypotheses

3.2.1 Study Objective 1

To assess if alcohol intake is associated with the incident cataracts among adults over a 17-year follow-up

Hypothesis: Alcohol use increases the risk of cataracts.

3.2.2 Study Objective 2

To assess if fruit and vegetable consumption is associated with the incident cataracts among adults over a 9-year follow-up

Hypothesis: Fruit and vegetable consumption reduces the risk of cataracts.

Chapter 4 Methods

This chapter describes the methodology used in the current study. Section 4.1 summarizes the study design and setting. Section 4.2 describes the data source. Section 4.3 provides a description of the study population, including the current study's inclusion and exclusion criteria. Section 4.4 details the variables in the analysis and how they are coded in our models. Section 4.4 illustrates the statistical methods used in our study. Section 4.5 presents other statistical considerations, such as handling tied data and missing data, sensitivity analysis and additional analysis.

4.1 Study Design and Setting

We conducted a population-based retrospective cohort study using a representative sample of Canadian residents aged 40 years or older selected from the Household component of the national population health survey (NPHS). We have two study objectives. The first study objective was to assess if alcohol intake is associated with the incident cataracts. Thus, the exposure of interest was alcohol intake for the first study objective. Our second study objective was to assess if fruit and vegetable consumption is associated with the incident cataracts. Thus, fruit and vegetable intake was the exposure of interest for the second study objective. Note that cataracts are the study outcome for both objectives.

Our study was conducted at the Research Data Centre at Western University. A research proposal for access to longitudinal master files was approved, so that information that is not available for the public could be obtained via Statistics Canada Research Data Centre at the Western University site (Appendix C).

We followed reporting guidelines for Strengthening the reporting of observational studies in epidemiology (STROBE) statement (Appendix D) (132).

4.2 Data Source

The current study utilized data from the national population health survey (NPHS) conducted by Statistics Canada starting in 1994. The survey collected health information and related socio-demographic information of the Canadian population. It included three components: The Household component was commenced in 1994 and carried out every two years through 18

years. The target population of the longitudinal NPHS Household component included residents from ten provinces in 1994/1995. Those living on Indian Reserves and Crown Lands, residents of health institutions, full-time members of the Canadian Forces Bases and some remote areas in Ontario and Quebec were excluded. The original sample of 17,276 was drawn from the Labour Force Survey in all provinces except Québec, where Enquête sociale et de santé was used. A multistage stratified systematic sampling method was applied in this survey. In the first stage, the geographic and socio-economic strata were formed within each province. In the second stage, six clusters were produced in each stratum using the probability proportional to size sampling method. In the next stage, the sample of dwellings was selected from clusters. Finally, the NPHS sample was obtained from the randomly selected households, and one family member was chosen as a respondent within each household (133).

A total of 17,276 respondents were included in Cycle 1, and the sample population was not renewed over time. The response rate for subsequent cycles was based on the original 17,276 individuals. The response rate ranged from 92.8% in Cycle 2 to 69.7% in Cycle 9. The data was collected via a Computer-Assisted Interview system. A test for the Computer-Assisted Interview system was conducted beforehand to identify any possible errors. Multiple approaches were adopted to reduce the non-response rate (i.e. interviewer training, use of multiple languages for interviews, non-respondent's follow-up, etc.). For example, an additional follow-up was conducted in April of the second year of each wave with those did not respond in the first year. Dates and causes of death were ascertained against the Canadian Vital Statistics Database by Statistics Canada.

4.3 Study Population

For the current study, the information regarding fruit and vegetable consumption was only collected from cycle 5 (2002/2003), while assessment of alcohol consumption was available from cycle 1 (1994/1995) through cycle 9 (2010/2011). Data from cycle 1 through cycle 9 were used for the study objective 1, and data from cycle 5 through cycle 9 for the study objective 2.

For study objective 1, participants aged 40 years and older in cycle 1 were included in the study cohort. We also added subjects into the study cohort as they became 40 years of age from cycle 2 (1996/1997) through cycle 8 (2008/2009). We excluded subjects who reported cataracts at the

time of cohort entry, which was done to ensure that our population was at risk at the beginning of follow-up. We also excluded those who had reported a history of steroid medication use at the cohort entry time.

Similar inclusion and exclusion criteria were applied for our study objective 2. Participants who were 40 years and older and did not report cataracts in cycle 5 were included in our initial cohort. Those who became 40 years and did not a history of cataracts during the time of cohort entry were added from cycle 6 to cycle 8. People with a history of steroid use were excluded from our study.

People under 40 years old were excluded because they were less likely to develop an age-related cataract. The inclusion of only those who did not report cataracts at enrollment of the cohort allows us to predict future incident cataracts to strengthen temporality. After this inclusion and exclusion were applied, the cohort for study objective 1 had a sample size of 9,899, and the cohort for study objective 2 had a sample size of 7,388. Figure 3 and Figure 4 demonstrates a flow chart of two cohorts' creation based on the study inclusion and exclusion criteria.

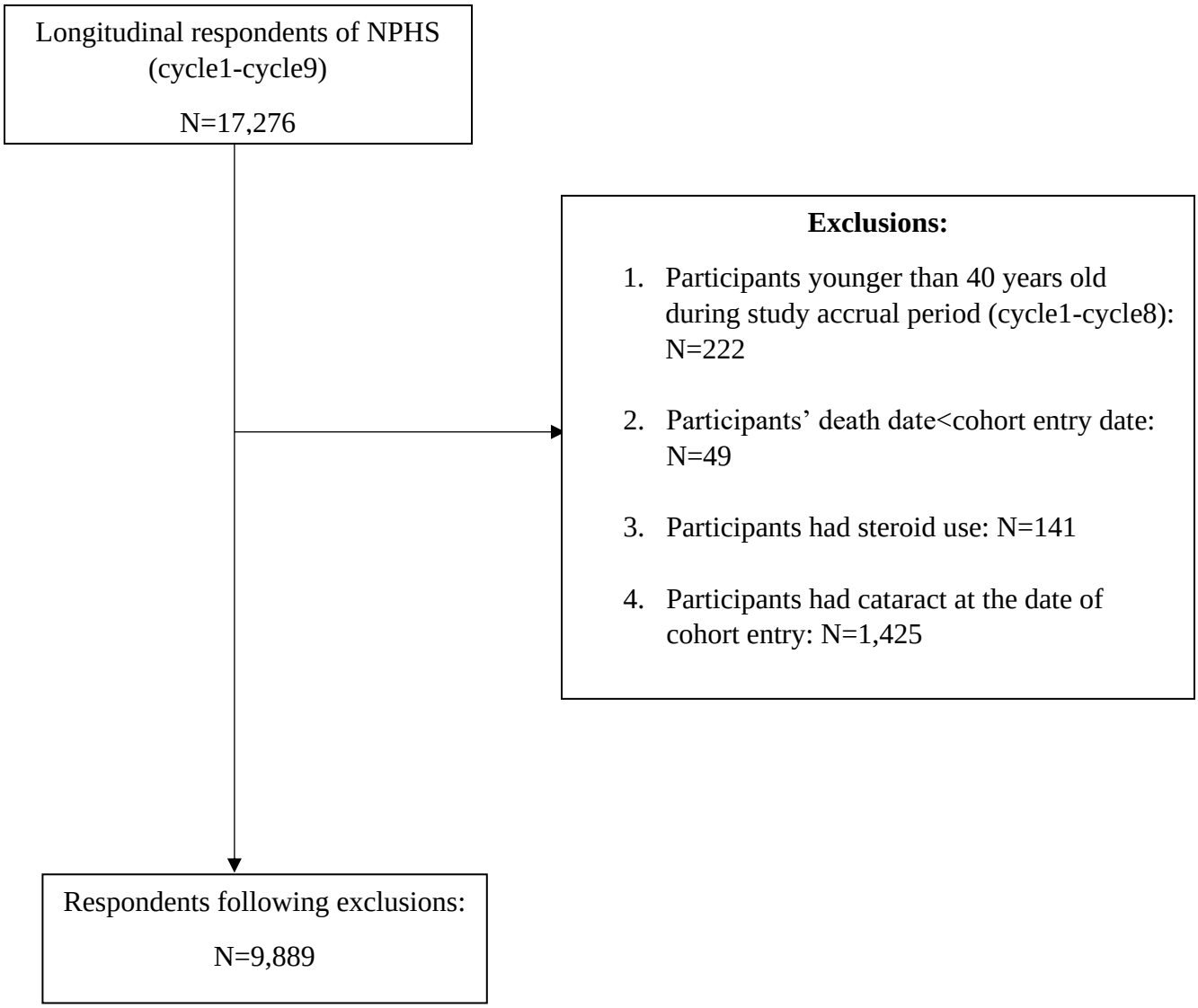


Figure 3. Study population flow chart for study objective 1

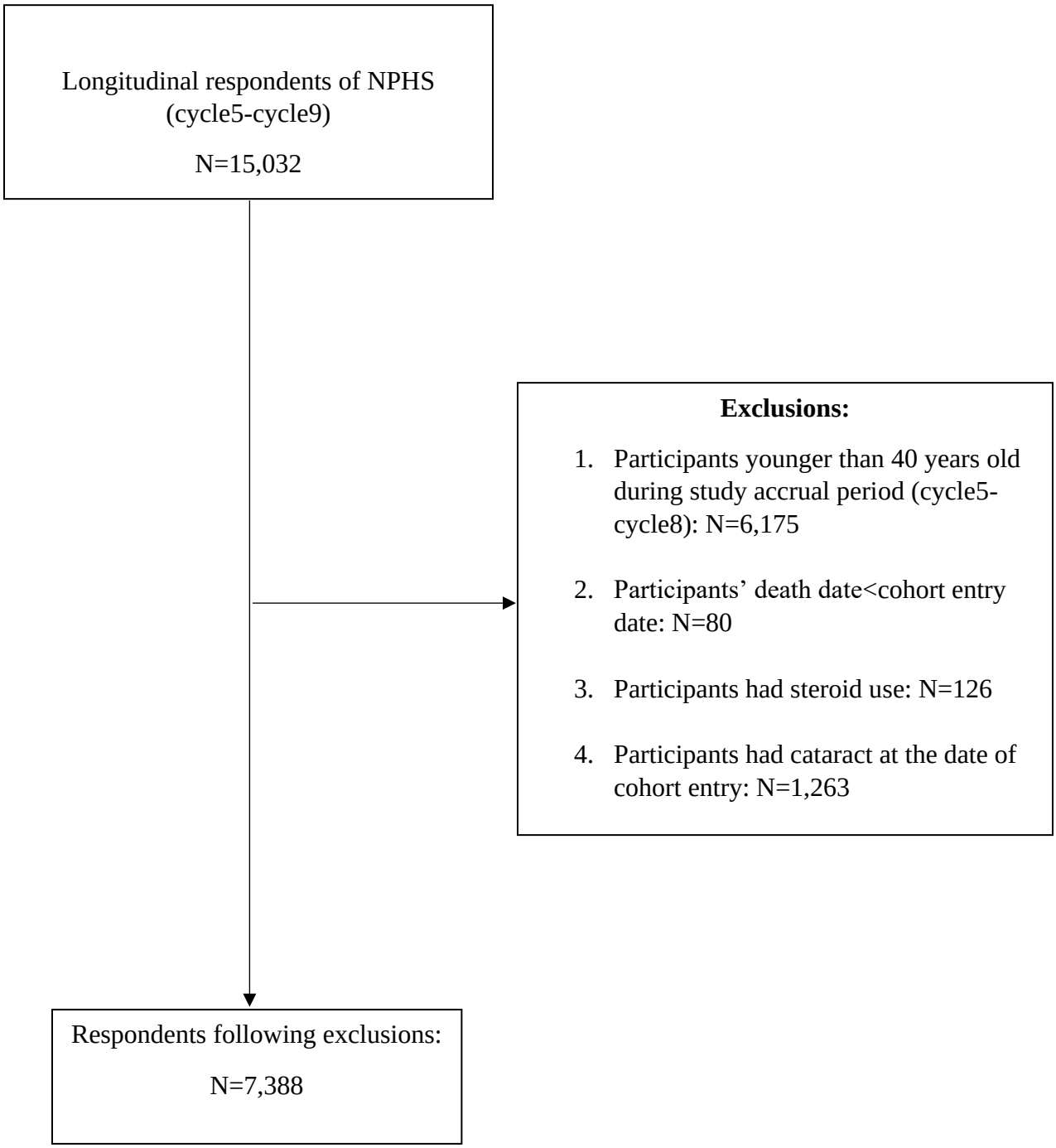


Figure 4. Study population flow chart for study objective 2

4.4 Variables

This section details the assessment of variables that were included in our data analysis.

4.4.1 Outcome

The cataract is our outcome of interest in this study. This variable was measured by asking respondents aged 18 years and older if they were ever diagnosed with cataracts by a health professional. The same question was repeatedly asked at each cycle. Values from those who answered: “don’t know”, “not applicable”, or “refusal” were coded as missing values. The incident cataract was identified by the initial report of a cataract. The time origin was the time of cohort entry. The observation was terminated by outcome-event occurrence, death, the end of the NHPS follow-up (2010/2011) or loss to follow up, whichever came first.

4.4.2 Exposures

4.4.2.1 Alcohol Intake

Alcohol intake is the primary exposure variable for our first study objective in the current study. Each respondent aged 12 and over was asked if they consumed alcohol (beer, wine, liquor or other alcoholic beverage) during the past 12 months. The frequency of alcohol use was determined by a question: “During the past 12 months, how often did you drink the alcoholic beverage?” Response categories were every day, 4 to 6 times a week, 2 to 3 times a week, once a week, 2 to 3 times a month, once a month and less than once a month. The quantity of alcohol use was ascertained by the amount of drinks consumed during the past week. Average daily alcohol consumption was derived by NPHS, which was equal to the weekly total alcohol use divided by 7. A drink was defined as one bottle or can of beer, one glass of wine or wine cooler, or one drink or cocktail with 1 and ½ ounces of liquor. We computed monthly alcohol intake by the following equation:

$$\textit{Monthly alcohol intake} = \textit{Frequency of monthly alcohol use} \times \textit{Average daily alcohol use}$$

where frequency of monthly alcohol use was averaged and converted from response categories of the frequency of alcohol use as following values: 30 (every day), 20 (4 to 6 times a week), 10

(2 to 3 times a week), 4 (once a week), 2.5 (2 to 3 times a month), 1 (once a month) and 0 (less than once a month), and average daily alcohol use was equal to the total alcohol intake for the past one week prior to interview divided by 7 (134).

4.4.2.2 Fruit and Vegetable Consumption

Fruit and vegetable (FV) consumption was the exposure variable for our second study objective. Information regarding fruit and vegetable use was obtained by asking the frequency of fruit juices (orange, grapefruit, or tomato etc.), fruit, green salad, potatoes, carrots, and other vegetables, respectively. Study subjects were asked to specify the reporting period, with response categories being daily, weekly, monthly, and yearly. We used the total daily consumption of fruit and vegetables in our study, which is an aggregate of daily FV intake computing by the following equation:

Daily FV consumption=Sum of annual FV consumption for each food item ÷ 365

It should be noted that FV intake reflects how frequently respondents were consumed during the recall period rather than an amount. This variable was kept as a continuous variable in our models (134).

4.4.3 Covariates

Based on a comprehensive literature review, two directed acyclic graphs (DAG) were drawn accordingly to establish a causal framework of this study (Figure 5 and Figure 6). The following variables were identified as confounders of the association of alcohol use with cataracts: age, sex, income, race, education, smoking, BMI, physical activity, use of vitamin/minerals, FV intake, UV exposure, diabetes and hypertension. The following variables were identified as confounders of the association of FV intake with cataract: age, sex, income, race, education, smoking, BMI, physical activity, use of vitamin/minerals, alcohol use, UV exposure, diabetes, and hypertension.

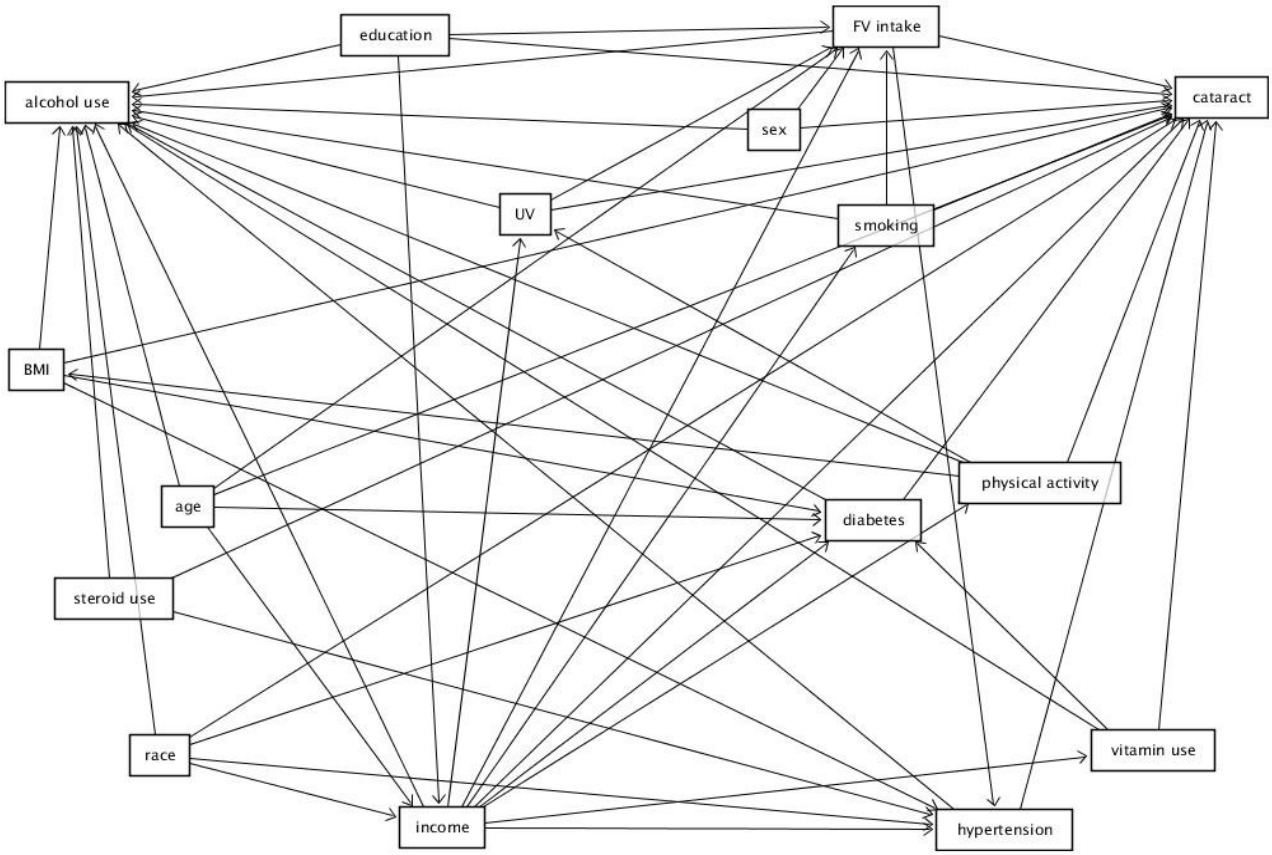


Figure 5. Directed acyclic graph (DAG) for examining the effects of alcohol use on cataracts with confounders.

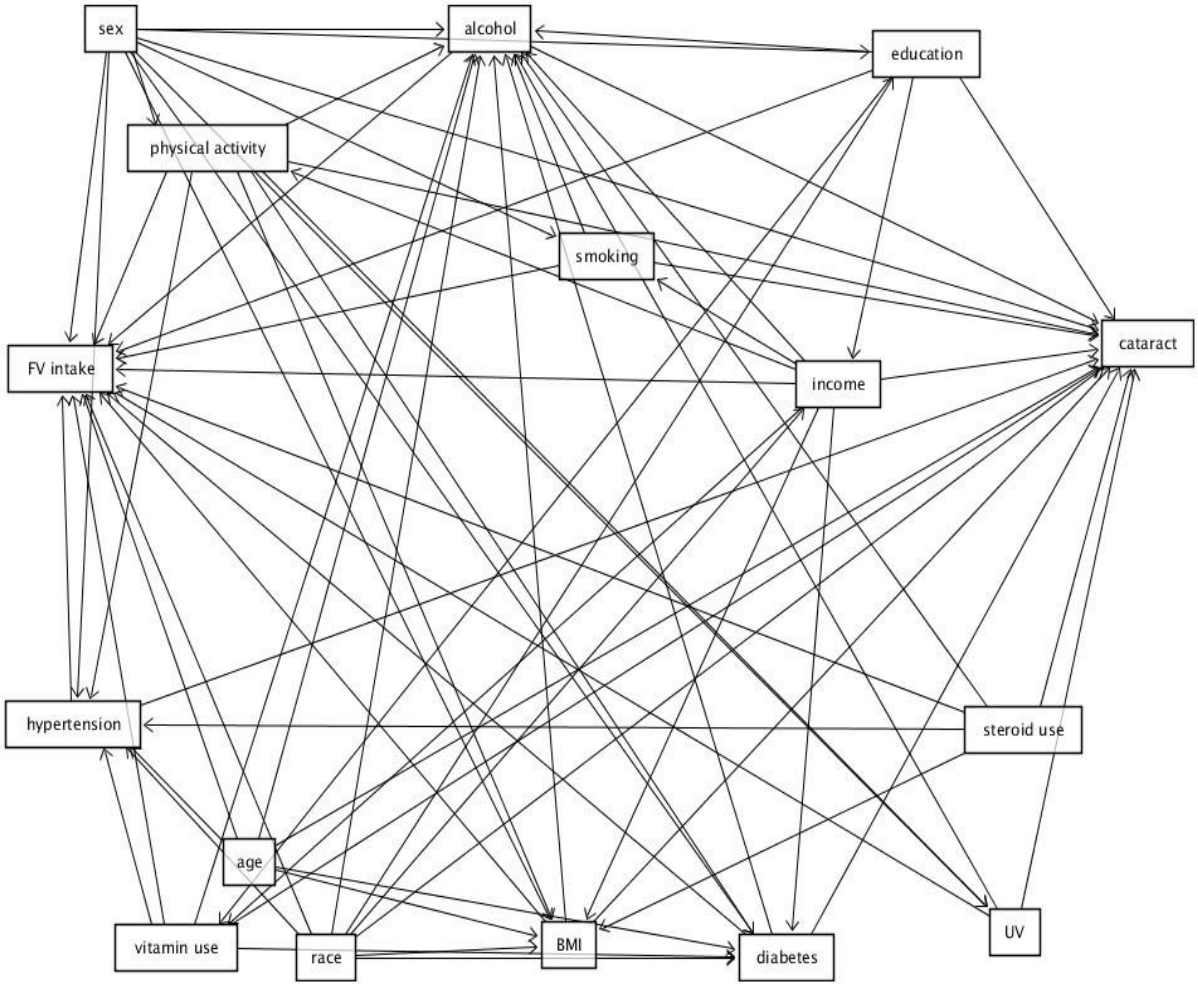


Figure 6. Directed acyclic graph (DAG) for examining the effects of FV intake on cataracts with confounders.

4.4.3.1 Age

NHPS reports every respondents' age in years by asking their date of birth and age. We calculated each respondent's age based on the year of the cohort entry and their birth year. In the current study, age was coded as a continuous variable, and a quadratic term in age was added into the model to address the possible non-linearity of the relation being modelled.

4.4.3.2 Sex

Sex (male or female) was ascertained in the NHPS survey. For our analysis, sex was kept as a binary variable, with male being the reference category.

4.4.3.3 Income

Income, a derived variable by NPHS, was included in the analysis. The income was divided into four categories based on the total annual household income and the number of people living in the household: lowest income, lower middle income, upper middle income, and highest income.

Lowest income was defined as: less than $15,000 and 1 or 2 persons; less than $20,000 and 3 or 4 persons; less than $30,000 and 5 or more persons. Lower middle income denoted that: $15,000 to $29,999 and 1 or 2 persons; $20,000 to $39,999 and 3 or 4 persons; $30,000 to $59,999 and 5 or more persons. Upper middle income represented: $30,000 to $59,999 and 1 or 2 persons; $40,000 to $79,999 and 3 or 4 persons; $60,000 to $79,999 and 5 or more persons. Highest income was defined as: $60,000 or more and 1 or 2 persons; $80,000 or more and 3 persons or more. Income was coded as a categorical variable.

4.4.3.4 Education

The highest level of educational attainment was reported in NHPS. The response was categorized into four classes: less than secondary school graduation, secondary school graduation, some post-secondary and post-secondary graduation. The education variable was treated as a nominal scale categorical variable.

4.4.3.5 Race

Race was determined by a question: “How would you best describe your race or colour?”. For this study, we regrouped race as White, Black, Asian, and other, with White being the reference category.

4.4.3.6 Smoking

Smoking was documented by asking a question: “At the present time, do you smoke cigarettes daily, occasionally or not at all?” The response was categorized as daily, occasionally, and not at all. Smoking status was coded as a binary variable, with ‘daily’ and ‘occasionally smoking’ treated as ‘smoker’ and ‘not at all’ as ‘non-smoker’.

4.4.3.7 BMI

The BMI was a derived variable available in NPHS. The formula for BMI is weight in kg, dividing by height in meters squared. Our model included it as a continuous variable.

4.4.3.8 Physical Activity

Physical activity (PA) was categorized into three categories: active, moderate, and inactive, which was based on the energy expenditure (EE) values during leisure time activities in the past three months. Energy expenditure values were computed by the following equation:

$$\boldsymbol{EE\ (kcal/kg\ /day) = Sum\ of\ ((Ni \times Di \times METS)/365)}$$

where Ni is the number of times respondents engaged in specific activity over 12 months, Di is the average duration in hours of the activity, and METS is the energy cost of the activity expressed as kilocalories expended per kilogram of body weight per hour of activity (kcal/kg per hour)/365 (to convert yearly data into daily data). Active is those who averaged 3.0+ kcal/kg/day of energy expenditure. Moderate is those who averaged 1.5 - 2.9 kcal/kg/day. Inactive is those with an energy expenditure below 1.5 kcal/kg/day. We coded physical activity as a categorical variable.

4.4.3.9 Use of Vitamin/Mineral Supplements

The use of vitamin/mineral supplements was examined by asking if the respondents had ever taken them. This variable was coded as a binary variable, with “yes” or “no”.

4.4.3.10 Hypertension

Assessment for hypertension status was performed by asking if subjects had this condition diagnosed by a health professional. The value of this variable was kept as “yes” or “no”.

4.4.3.11 Diabetes

Ascertainment for diabetes was performed by asking if the participants had this condition diagnosed by a health professional. Diabetes was coded as “yes” or “no”.

4.5 Statistical Analysis

4.5.1 Descriptive Statistics

Descriptive statistics were presented for the characteristics of the study population. Frequency distributions were provided for categorical variables, including sex, age groups, years of cohort entry, alcohol use categories, FV intake groups, income, education, race, smoking, physical activity, use of vitamin/minerals supplements, hypertension, and diabetes. For the continuous variables, including age, BMI, alcohol use and FV consumption, mean value, standard deviation, median value and interquartile range were provided. Next, the same descriptive statistics were depicted for each category of alcohol intake and fruit and vegetable consumption.

In addition, numbers of incident cases and cumulative incidence of cataract occurrence using cumulative incidence function over various follow-up time periods were presented (135). Finally, the incidence rate per 1,000 person-years was reported.

4.5.2 Cox Proportional Hazards Analysis for Study Objective 1

The first objective of this study was to examine the effect of alcohol intake on incident cataracts.

To address the study objectives, Cox proportional hazards model was fitted in our study. The benefit of Cox proportional hazards model is that it is a semiparametric model, and baseline hazard need not be specified.

Both crude and adjusted hazard ratio (HR) was estimated through the Cox proportional hazards model. The crude model examined how alcohol use affects cataracts without including any covariates. The adjusted model was controlled for age, sex, year of cohort entry, income, education, race, physical activity, smoking, BMI, vitamin/mineral supplements, hypertension, and diabetes. A Wald test was used to obtain confidence interval. A two-sided p value <0.05 was considered statistically significant.

One of the Cox proportional hazards model's essential features is that it allows time-varying covariates to be included in the model. In our study, some variables can change over time, i.e. alcohol intake and BMI. Estimation could be distorted if these variables were only measured at baseline and included in the model as time-independent variables (136,137). As cataracts status and time-dependent covariates were updated at each wave, the dataset was lagged one year so that these explanatory variables at time (t-1) predicted future incident cataracts to avoid reverse causality. We hypothesize that cataracts are associated with the cumulative use of alcohol. Thus, we incorporated alcohol use as a cumulative intake and updated it at each cycle. The specifications of the fitted proportional hazards Cox models are shown below:

$$\log\left(\frac{h_i(t)}{\alpha(t)}\right) = \beta_1 x_{i1}(t-1) + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \beta_8 x_{i8} + \beta_9 x_{i9} + \beta_{10} x_{i10} + \beta_{11} x_{i11} + \beta_{12} x_{i12} + \beta_{13} x_{i13} + \beta_{14} x_{i14} + \beta_{15} x_{i15} + \beta_{16} x_{i16}(t-1) + \beta_{17} x_{i17} + \beta_{18} x_{i18} + \beta_{19} x_{i19} + \beta_{20} x_{i20}$$

where *t* denotes time in years and *i* represents *ith* observation.

$h_i(t)$=hazard at time t

$\alpha(t)$ = basline hazard function

$x_1(t-1)$=cumulative alcohol use up to time (t-1) (10 drinks/month)

x_2 =age (years)

x_3=sex (0=male and 1=female)

x_4 = lower middle income (0=lowest income, 1=lower middle income)

x_5 = upper middle income (0=lowest income, 1= upper middle income)

x_6 = highest income (0=lowest income, 1= highest income)

x_7 = secondary school graduation (less than secondary school graduation=0, secondary school graduation=1)

x_8 = secondary school graduation (less than secondary school graduation=0, secondary school graduation=1)

x_9 = post − secondary graduation(less than secondary school graduation=0, post-secondary graduation =1)

x_{10} = inactive (active=0, inactive =1)

x_{11} = moderately active (active=0, moderately active =1)

x_{12} = Asian race(white=0, Asian =1)

x_{13} = Black race(white=0, Black =1)

x_{14} = Other race(white=0, other =1)

x_{15} = smoking (yes=1 and no=0)

$x_{16}(t-1)$ = BMI at time (t-1) (kg/m^2)

x_{17} = year of cohort entry (1994-2009)

x_{18} = vitamin use (yes=1 and no=0)

x_{19} = hypertension (yes=1 and no=0)

x_{20} = diabetes (yes=1 and no=0)

$\beta_1, \beta_2 \ldots\ldots \beta_{20}$ are the corresponding regression coefficients. We used cumulative alcohol intake in our analysis, which accumulated alcohol intake from all cycles (including information prior to the date of cohort entry). The equation to compute cumulative alcohol intake was proposed by Allison (138):

$$\boldsymbol{Cum(t)=(Cum(t\text{-}1) \times (t\text{-}1) + al(t))/t}$$

where *Cum(t)* represents cumulative alcohol intake by time *t*, *t* is time in years, *Cum(t-1)* denotes cumulative alcohol intake by time (t-1), and *al(t)* is average monthly alcohol intake at time t (drinks/month).

4.5.3 Cox Proportional Hazards Analysis for Study Objective 2

The second objective of this study was to examine the effect of fruit and vegetable (FV) intake on incident cataract.

Since fruit and vegetable information was only collected from cycle 5, the data from cycle 5 through cycle 9 was used to address this study objective. Both crude and adjusted hazard ratio (HR) was estimated through the Cox proportional hazards model. The crude model examined how FV intake affects cataracts without including any covariates. The adjusted model was controlled for age, sex, income, education, race, physical activity, smoking, alcohol use, BMI, year of cohort entry, use of vitamin/minerals supplements, hypertension, and diabetes. A Wald test was used to obtain confidence interval. A two-sided p value <0.05 was considered statistically significant.

The fitted Cox proportional hazards models are shown below:

$$\log\left(\frac{h_i(t)}{\alpha(t)}\right) = \beta_1 x_{i1}(t-1) + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \beta_8 x_{i8} + \beta_9 x_{i9} + \beta_{10} x_{i10} + \beta_{11} x_{i11} + \beta_{12} x_{i12} + \beta_{13} x_{i13} + \beta_{14} x_{i14} + \beta_{15} x_{i15} + \beta_{16} x_{i16}(t-1) + \beta_{17} x_{i17} + \beta_{18} x_{i18} + \beta_{19} x_{i19} + \beta_{20} x_{i20} + \beta_{21} x_{i21}(t-1)$$

where *t* denotes time in years and *i* represents *ith* observation.

$h_i(t)$=hazard at time t

$\alpha(t) =$ basline hazard function

$x_1(t-1)$=cumulative alcohol use up to time (t-1) (10 drinks/month)

x_2 =age (years)

x_3=sex (0=male and 1=female)

x_4 = lower middle income (0=lowest income, 1=lower middle income)

x_5 = upper middle income (0=lowest income, 1= upper middle income)

x_6 = highest income (0=lowest income, 1= highest income)

$x_7 =$ secondary school graduation (less than secondary school graduation=0, secondary school graduation=1)

$x_8 =$ secondary school graduation (less than secondary school graduation=0, secondary school

graduation=1)

x_9 = post − secondary graduation(less than secondary school graduation=0, post-secondary graduation =1)

x_{10} = inactive (active=0, inactive =1)

x_{11} = moderately active(active=0, moderately active =1)

x_{12} = Asian race(white=0, Asian =1)

x_{13} = Black race(white=0, Black =1)

x_{14} = Other race(white=0, other =1)

x_{15} = smoking (yes=1 and no=0)

$x_{16}(t-1)$ = BMI at time (t-1) (kg/m^2)

x_{17} = year of cohort entry (1994-2009)

x_{18} = vitamin use (yes=1 and no=0)

x_{19} = hypertension (yes=1 and no=0)

x_{20} = diabetes (yes=1 and no=0)

$x_{21}(t-1)$ = cumulative FV intake up to time (t-1) (servings/day)

$\beta_1, \beta_2 \ldots\ldots \beta_{21}$ are the corresponding regression coefficients. We used cumulative alcohol intake in our analysis, which accumulated alcohol intake from all cycles (including information prior to the date of cohort entry). The equation to compute cumulative alcohol intake was proposed by Allison (138):

$$\boldsymbol{Cum(t)=(Cum(t\text{-}1) \times (t\text{-}1) + al(t))/t}$$

where *Cum(t)* represents cumulative alcohol intake by time *t*, *t* is time in years, *Cum(t-1)* denotes cumulative alcohol intake by time (t-1), and *al(t)* is average monthly alcohol intake at time t (drinks/month).

4.6 Other Statistical Considerations

4.6.1 Censoring

Three sources of censoring might arise in the current study:1) termination of the study.2) loss to follow-up and 3) death. Response pattern, a variable in NPHS, was used to determine response status of each participant throughout the cycles. If the loss to follow-up occurred at least in two

successive cycles, the subject was considered to be right-censored. The time of censoring was the last time of interview. We specified this because some participants might not respond in one cycle but could come back for an interview in the next cycle. The NPHS was already linked to the Canadian Vital Statistics Death Database to identify the date of death for each participant.

The main assumption underlying survival analysis is non-informative censoring: subjects that are censored have the same probability of experiencing an event of interest as individuals that remain in the study (135). In the primary analysis, we censored subjects at the time of death, which might violate the assumption of non-informative censoring (139). We therefore conducted an additional analysis that consider death as a competing risk (140).

4.6.2 Missing Data

Response categories of "don't know", "refuse", "not applicable", or "not stated" were coded as missing data in our study. Observations with one or more missing data are excluded from the analysis in complete case analysis. Although the percentage of missing data for one variable might be small, the combination of the variables with missing value could constitute a relatively large percent of a dataset. As a consequence, the complete case analysis could substantially reduce the sample size and result in biased estimates.

Multiple imputations were utilised in our sensitivity analysis to examine the robustness of our complete case analysis. The percentage of each missing item and the missing pattern was checked before performing multiple imputations. Missing data in our study arose from response categories of "don't know", "refuse", "not applicable", or "not stated". Missing data were distributed across variables in a an arbitrary pattern , which was illustrated in Appendix F and Appendix G (141). We assumed the missing mechanism in our data was missing at random (MAR) .We think this assumption may be reasonable because the missingness can be explained by observed values (142). For example, a recent study showed that age and lower education level were associated with the non-response rates in self-reported health survey (143). The method we used was the fully conditional method which uses a separate conditional distribution for each imputed variable, as our data was a mixed set of continuous, nominal, ordinal and binary variables. (144,145). We used linear regression to impute continuous variables, ordinal logistic regression for binary and ordinal variables, and discriminant function for nominal data.

The variables we used in the imputation model included all the dependent, independent variables and their quadratic/interaction terms contained in the previous regression models. Additionally, we included Nelson-Aalen estimator of hazard function and event indicator in our imputation model (146). We performed 30 imputations and used PROC MIANALYZE to estimate pooled parameters.

4.6.3 Tied Data

Cataracts were assessed every two years in our dataset. As a result, tied event times were included in our data. Efron's approximation method was used in our analysis to address issues of tied events times, which has been proved to be efficient and valid to handle tied data, even with heavy tied data (147,148).

4.6.4 Additional Analysis

Several additional analyses were performed in our study. First, age was explored for its interaction effect on the association between alcohol use and cataracts, followed by an effect modification analysis of sex. The interaction effect of age and alcohol intake on the cataracts was examined by including an interaction term in our model. The same method was also applied to assess the interaction effect of sex. Similar analyses were conducted to estimate the interaction effect of age and sex with fruit and vegetable intake. Further subgroup analyses were carried out by stratifying the age ($\leq$ 65 versus >65 years) and sex (male versus female).

Additionally, we treated death as a competing risk in our analysis, as death could preclude the occurrence of our event of interest. Fine and Gray's sub-distribution hazard model was fitted to estimate the effects of covariates on the outcome of interest in the presence of competing risk (140).

4.6.5 Sensitivity Analysis

Sensitivity analysis was performed for the current study to assess the robustness of estimation. First, the sensitivity analysis was conducted to address the unmeasured confounders. Since the information regarding fruit and vegetable consumption was only collected from cycle 5 in NPHS, it could not be controlled for our study objective 1. Additional models with and without

additionally controlling fruit and vegetable intake were fitted using data from cycle 5 through 9 to examine the robustness of analysis for study objective 1.

Second, we also conducted analyses that excluded outliers of alcohol intake (>180 drinks/month) and FV consumption (>10 servings/day) in our model to examine the influence of outliers on our estimate.

Finally, we limited our study population to white respondents to control the possible confounding effect of race.

4.6.6 Assessment of the Proportional Hazards Assumption

One key assumption of Cox model is proportional hazards, which assumes that the hazard ratio of interest is constant over time. It is essential to try to detect the proportional hazards for covariates of interest. The proportional hazards assumption was evaluated by Schoenfeld test, which examines the scaled Schoenfeld residuals and time (149). The results showed that Schoenfeld residuals of either exposures of interest was not statistically significantly associated with three functions of time, indicating that the proportional hazards assumption seemed to hold for our exposures (Appendix W and Appendix X).

4.7 Statistical Analysis Software

SAS version 9.4 (SAS Institute Inc. 2012) was used for all statistical analysis. Descriptive statistics were produced using PROC FREQ and PROC MEANS procedures. Cox proportional hazards regression analyses were performed using PROC PHREG procedures. Multiple imputations were conducted using PROC MI procedure and PROC MIANALYZE was used for pooled parameter estimation.

Chapter 5 Results

This chapter provides the results of this study. Section 5.1 provides descriptive statistics of the study population. Section 5.2 presents the cumulative incidence and incidence rate of cataracts in the study population.

5.1 Descriptive Statistics

5.1.1 Overall Study Population Characteristics

After exclusions, we identified 9,889 respondents over the ages of 40 between 1994 and 2009 from NPHS (Figure 3). A total of 7,388 respondents aged 40 years or older were included from 2002 to 2009 (Figure 4). Baseline characteristics of included participants for each cohort were reported in Table 1 and Table 3, respectively.

The mean age of members of our first cohort at time of cohort entry was 52 years, and the median was 48 years. Approximately 55% of the respondents in our cohort were female. Most of the respondents (69.9%) entered the cohort at cycle1 (1994-1995). Approximately a third of participants received the highest educational attainment (32.9%). More than half of the subjects had middle income, with 26.2% of lower middle income and 32.9% of upper middle income. 25.2% of respondents reported smoking in the past year of cohort entry. White respondents constituted the majority our study population (86.9%). The mean BMI of the sample population was 26.2kg/m^2. The majority of the members of our cohort were physically inactive (57.1%). In terms of chronic diseases, only a small percentage of subjects had diabetes (4.2%), while 14.0% of participants had hypertension. Over one-third of respondents (35.0%) reported vitamins/minerals use in the reporting period. Approximately two-thirds (66.0%) of subjects did not report alcohol intake, while only 2.8% consumed more than 60 drinks/month of alcohol. The missing data for most variables were relatively low (≤ 5%), while 7.4% of the sample population did not report smoking, and 8.2% of respondents had a missing value for the race.

The mean age of our second cohort at time of cohort entry was 55 years old, and 55.3% were women. A total of 6,180 subjects (83.7%) were included into the study at the beginning of cohort entry (cycle 5). Over one-third of participants had post-secondary education. The upper middle-income category constituted the largest portion of the population (31.4%), compared with other

income categories. Similarly, white subjects made up around 90% of the study population. 22.3% of respondents reported they had a smoking history in the past year. Over half of the study population did not report alcohol use in the reporting period, whereas only 1.3% reported consuming more than 20 drinks per month. The distribution of BMI was also similar to the first cohort, with a mean value of 26.8 kg/m^2. Over half of the sample population (50.2%) were physically inactive. Hypertension was the most common chronic disease (21.5%), while only 6.7% of subjects had diabetes. Our study found that over 40% of the population took vitamins or supplements, and most respondents consumed 3 to 6 servings of fruit and vegetable per day.

5.1.2 Baseline Characteristics across Alcohol Use and FV Consumption Categories

Baseline characteristics across alcohol intake and FV categories were presented in Table 2 and Table 4.

From Table 2, alcohol users were more likely to be older. The mean age of the 20-59.9 years and ≥ 60 years groups were 54.5 years and 58.3 years, respectively, while it was 42.7 years in the 0.1-9.9 (drinks/month) group. Compared with men, women were less likely to drink alcohol, with only 17.6% of female consuming more than 60 drinks per month, while females constituted 55.2% of non-drinkers. There was a tendency that increased alcohol use was associated with lower educational attainment. The percentage of less than secondary graduation rose from 10.8% for 0.1-9.9 (drinks/month) alcohol use group to 31.9% for ≥ 60 (drinks/month) group. Alcohol use was associated with income as well. The low-income per cent doubled from 6.0% in 0.1-9.9 (drinks/month) alcohol use group to 12.2% in ≥ 60 (drinks/month) group, while the percentage of highest income decreased from 32.6% to 17.9% in these groups. The smoking percentage was higher as individuals consume more alcohol, increasing from 29.2% to 39.1%. The percentage of both hypertension and diabetes rose from 5.1% to 19.0% and 1.5% to 6.1%, respectively, with the increased use of alcohol. BMI, race, and physical activity remained similar across different alcohol use categories. The proportion of vitamin use was quite similar throughout the alcohol use groups, except for the ≥ 60 (drinks/month) groups, with a lower percentage (25.4%).

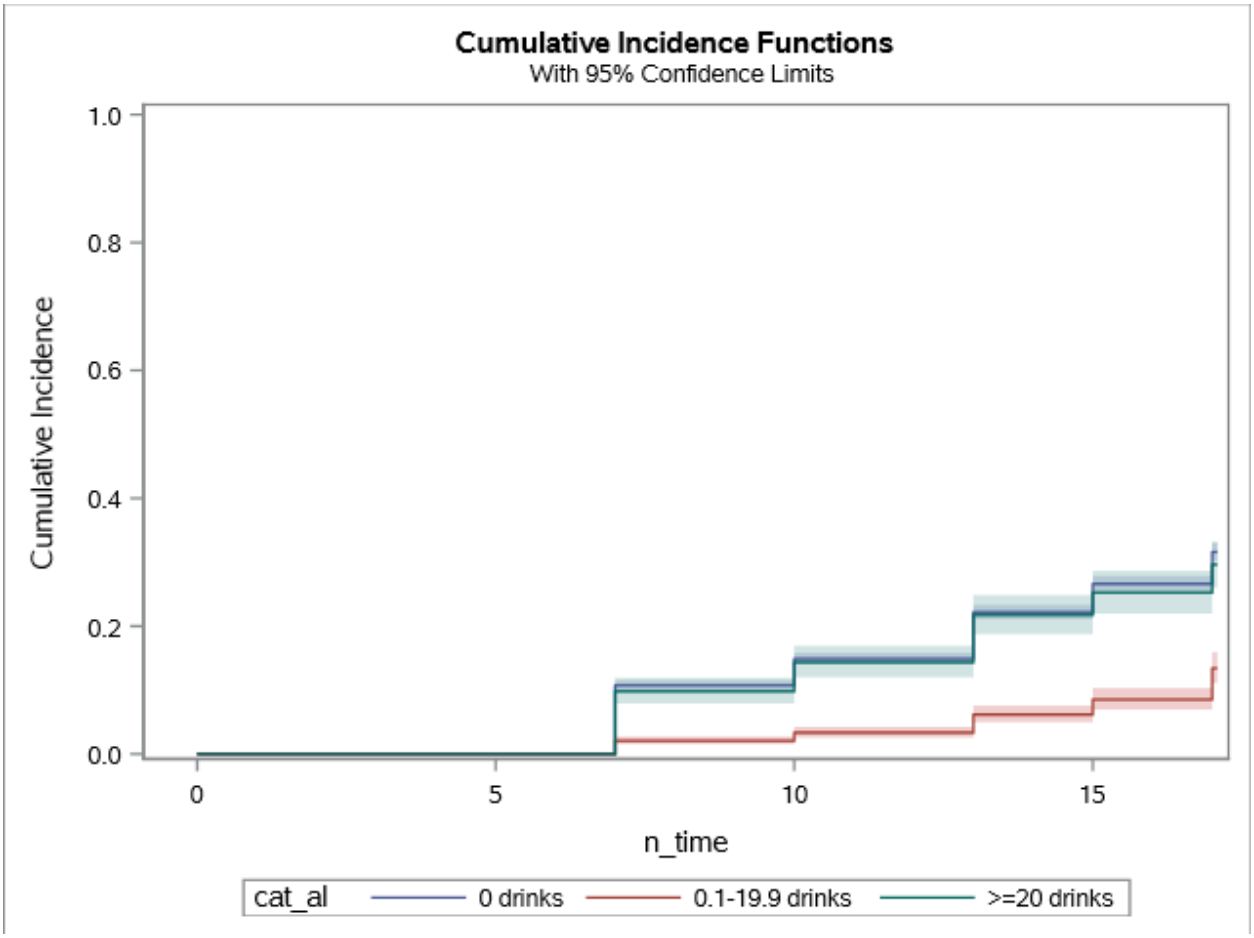
Table 4 presents baseline characteristic across fruit and vegetable (FV) intake categories. Mean age was slightly higher as individuals consumed more FV (51.9 years to 54.5 years). More

females (63.0%) were observed in the higher FV consumption group. The people with higher education were more likely to consume FV, and the opposite trend was observed among those with lower education. For example, the frequency of higher education attainment increased from 30.0% for lower FV consumption to 40.0% for higher FV consumption. A similar effect was also shown for income categories. The percentages of the first two income categories decreased as participants ate more FV, and the opposite was true for the higher income categories. The prevalence of physical inactivity was higher among those eating less FV (57.9% vs 36.7%). The distributions of race, diabetes, hypertension, and BMI were quite similar across FV categories. Vitamin use was the highest in the ≥6 servings group (44.1%).

5.2 Cumulative Incidence at Different Follow-up Time and Overall Incidence Rate

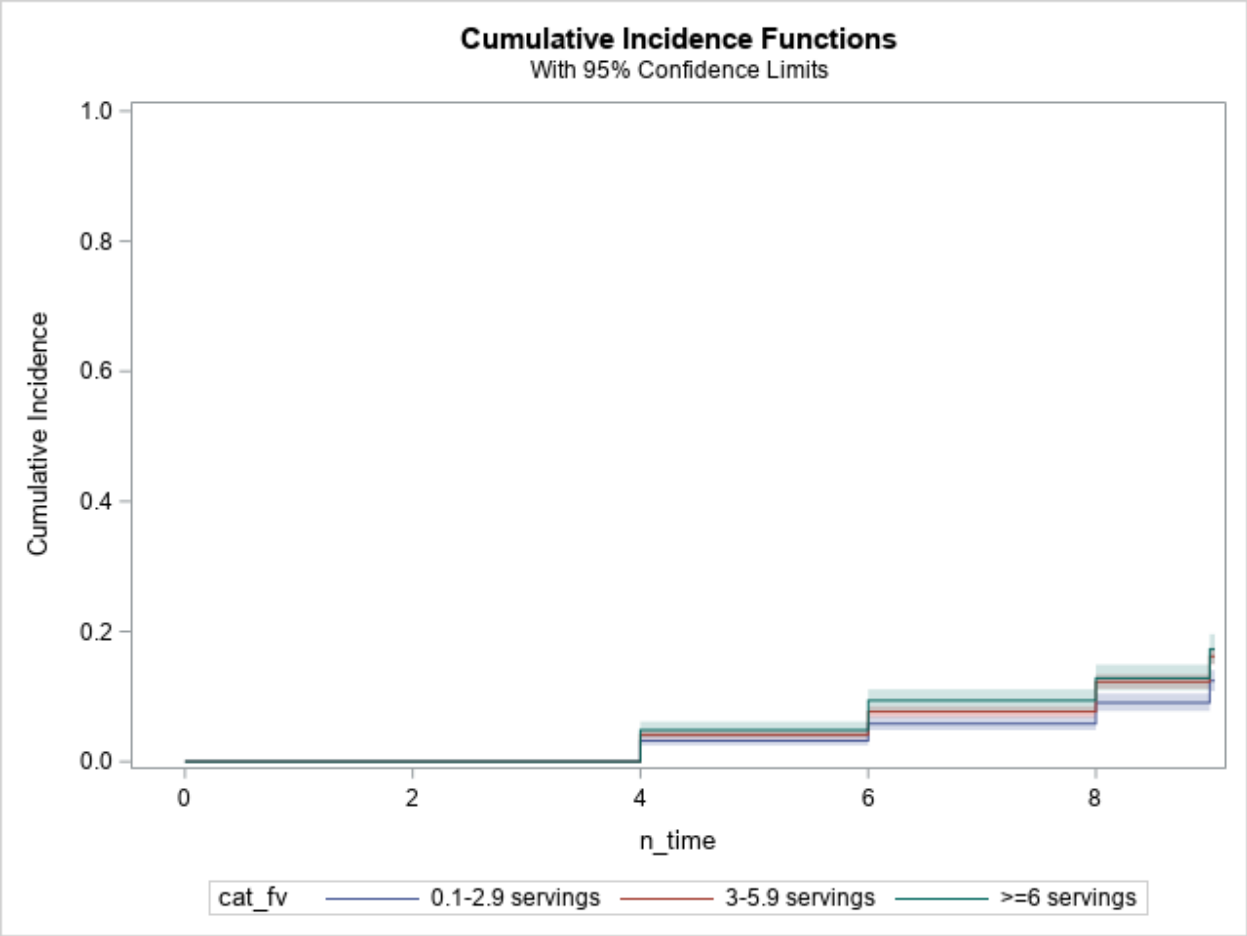
Figure 7 shows the cumulative incidence function of cataracts for NPHS respondents (1994-2011). The median follow-up time was 10 years, ranging from 1 to 17 years. The cumulative incidence of cataracts was 3.1% over two years of follow-up time and 30.3% over 17 years of follow-up. Over the study follow-up period from 1996/1997 to 2010/2011, 1,978 incident cataracts were identified, with an incidence rate of 19.2 cases per 1000 person-years (Appendix E).

Figure 8 presents the cumulative incidence function of cataracts for NPHS respondents (2002-2011). The follow-up time varied from 1 to 9 years, with a median of 8 years. Over the study follow-up period between 2004/2005 and 2010/2011, 1,099 incident cataracts diagnoses were reported. The cumulative incidence of cataract over nine years was 16.3%. The incidence rate was 19.7 cases per 1000 person-years (Appendix F).



n_time= follow-up time (years), cat_al= alcohol intake categories

Figure 7. Estimated cumulative incidence function of cataracts by alcohol intake categories for NPHS respondents (1994-2011).



n_time= follow-up time (years), cat_fv= fruit and vegetable consumption categories

Figure 8. Estimated cumulative incidence function of cataracts by fruit and vegetable consumption categories for NPHS respondents (2002-2011).

5.3 The Association of Alcohol Use and Cataracts

The results of fitting the crude and adjusted Cox proportional hazards models for the association between alcohol use and incident cataracts are summarized in Table 5. The results show that the crude hazard ratio was 1.01 (95% CI: 0.99 to 1.04; P=0.38). After the potential confounders were controlled for, the hazard of cataracts increased by 2% with every 10 drinks/month of alcohol use, although the association was not statistically significant (HR:1.02; 95% CI: 0.99 to 1.04; P=0.15). Table 6 shows the crude and adjusted models, including quadratic terms, the association was not statistically significant for both models (HR: 1.01; 95% CI: 0.99 to 1.04; P=0.38 and HR: 1.00; 95% CI: 0.95 to 1.04; P=0.83, respectively).

Table 7 presents the results of fitting the crude and adjusted Cox proportional hazards regression models, including time-varying covariates for the association of alcohol use with cataracts. The results showed that every 10 drinks of cumulative alcohol use per month were associated with a 3% increase of incident cataracts after controlling for the potential confounders (HR: 1.03; 95% CI: 0.98 to 1.09; P=0.25).

5.4 Sensitivity Analysis for Study Objective 1

In Table 8, parameter estimates using fully conditional specification (FCS) multiple imputations are shown. The missing data pattern is attached in Appendix U. The hazard ratio of alcohol use after multiple imputations shows a similar point estimate with a narrower confidence interval (HR: 1.02; 95% CI: 1.00 to 1.04; P=0.06) compared with the previous complete case analysis model (HR: 1.02; 95% CI: 0.99 to 1.04; P=0.38). Table 9 displays the multiple imputation models, including time-varying covariates. No statistically significant association was found (HR: 0.99; 95% CI: 0.90 to 1.08; P=0.83). The next two tables (Table 10 and Table 11) show the crude and adjusted Cox proportional hazards models when we excluded outlier of alcohol use (>180 drinks/month) or limited study population to white respondents. The adjusted HR were 1.00 (95% CI: 0.95 to 1.04; P=0.83) and 1.00 (95% CI: 0.95 to 1.05; P=0.95), respectively.

Lastly, we used data from cycle 5 through cycle 9 without additionally controlling for FV intake. The adjusted HR was 0.93 (95% CI: 0.64 to 1.34; P=0.69) (Table 19). Table 20 shows a similar model but contains time-varying covariates. Both models had slightly different point estimates

from models with additionally adjusting for FV intake, but the association was still statistically non-significant.

5.5 Additional Analysis for Study Objective 1

Appendix G shows the results of effect modification of age for the association between alcohol intake and cataracts. The interaction term was not statistically significant (P=0.95).

Appendix H displays the results for the interaction between alcohol use and sex on the effect of cataracts. We found no statistically significant interaction (P=0.34).

Appendix I and J show the subgroup analysis by age (> 65 years versus ≤ 65 years). The adjusted HR was 1.03 (95% CI: 0.95 to 1.11) and 1.01 (95% CI: 0.95 to 1.08), respectively.

Appendix K and L show the subgroup analysis by sex (female versus male). The adjusted HR was 1.10 (95% CI: 0.98 to 1.23) and 0.97 (95% CI: 0.91 to 1.03), respectively.

When death was treated as a competing risk using a Fine and Gray model, a similar result was observed for alcohol use, with an adjusted HR being 1.00 (95% CI: 0.94 to 1.06; P=0.87), compared to the Cox model with death as a censoring event (HR:1.00; 95% CI: 0.95 to 1.04; P=0.83) (Appendix M).

5.6 The Association of FV Consumption and Cataracts

Table 12 presents the crude and adjusted Cox proportional hazards model for the association between FV consumption and cataracts. The crude model showed that the hazard of cataracts increased by 6% (HR: 1.06; 95% CI: 1.04 to 1.09; P<0.0001) for every serving of FV per day. However, after adjusting for potential confounders, the estimated HR was 1.00 (95% CI: 0.96 to 1.05; P=1.00).

Table 13 describes the crude and adjusted Cox proportional hazards models, including quadratic terms of continuous variables to address the potential non-linearity effect. The results showed that the quadratic term of FV consumption was not statistically significant (P=0.47). The adjusted HR for FV intake was 1.04 (95% CI: 0.91 to 1.19; P=0.52)

Table 14 shows the crude and adjusted Cox proportional hazards model for the association of FV intake and cataracts, including time-varying covariates. The results showed that one daily serving of FV was associated with a 7% increase in cataracts (HR=1.07; 95% CI: 1.01 to 1.12; P=0.01). The HR was 1.03 after controlling for potential confounders (95% CI: 0.95 to 1.12; P=0.42)

5.7 Sensitivity Analysis for Study Objective 2

In Table 15, a model-based imputation using the fully conditional specification (FCS) method is shown. The missing data pattern is attached in Appendix V. The hazard ratio of FV intake after multiple imputations showed a similar point estimate with a narrower confidence interval (HR: 1.01; 95% CI: 0.98 to 1.04; P=0.51) compared with the previous complete case analysis. Table 16 displays the multiple imputation models, including time-varying covariates. No statistically significant association was found (HR: 0.99; 95% CI: 0.89 to 1.11; P=0.92). The next two tables (Table 17 and Table 18) show the results when we excluded outlier of FV consumption (>10 servings/day) and limited the study population to white respondents. The adjusted HR was 1.09 (95% CI: 0.88 to 1.34; P=0.43) and 1.05 (95% CI: 0.92 to 1.20; P=0.49), respectively.

5.8 Additional Analysis for Study Objective 2

Appendix N displays the results of effect modification of age on FV intake, followed by an interaction analysis between sex and FV consumption (Appendix O). The addition of FV-by-sex and FV-by-age was not statistically significant (P=0.48 and P=0.60, respectively), indicating no sex or age difference in the effect of FV on cataracts..

Appendix P and Q show the subgroup analysis by age (> 65 years versus ≤ 65 years). The adjusted HR was 1.00 (95% CI: 0.87 to 1.16) and 1.30 (95% CI: 0.99 to 1.71), respectively.

The results are stratified by sex (female versus male) in Appendix R and S. The adjusted HR was 1.04 (95% CI: 0.88 to 1.23) and 1.01 (95% CI: 0.79 to 1.30), respectively

Appendix T summarizes the results of subdistribution hazards using a Fine and Gray model treating death as a competing risk. The adjusted HR for FV intake was 1.04 (95% CI: 0.90 to 1.19; P=0.63), which was very similar to the results produced by the Cox proportional hazards model with death as a censoring event (HR: 1.04; 95% CI: 0.91 to 1.19; P=0.52).

Table 1. Baseline characteristics of participants aged 40 years or older from National Population Health Survey (1994-2009)

Characteristic*	Total Population (n=9,889)[a]
Age, years, mean (SD)	52.4 (13.1)
Age, years, median (IQR)	48 (21)
Age groups	
40-49	5342 (54.0%)
50-59	1697 (17.2%)
60+	2850 (28.8%)
Sex (Female)	5407 (54.7%)
Education[b]	
less than secondary education	2960 (29.9%)
secondary graduation	1341 (13.6%)
some posts-secondary	2279 (23.1%)
post-secondary	3254 (32.9%)
missing	55 (0.6%)
Income[c]	
lowest income	1665 (16.9%)
lower middle income	2595 (26.2%)
upper middle income	3254 (32.9%)
highest income	1844 (18.7%)
missing	531 (5.4%)
Race	
Asian	303 (3.1%)
Black	94 (1.0%)
White	8594 (86.9%)
Other	83 (0.8%)
missing	815 (8.2%)
Smoking (yes)	2554 (25.8%)
missing	736 (7.4%)
BMI (kg/m^2), mean(SD)	26.2 (4.6)
missing	173 (1.8%)

Physical activity[d]	
active	1544 (15.6%)
moderate	2166 (21.9%)
inactive	5648 (57.1%)
missing	531 (5.4%)
Use of vitamin (yes)	3456 (35.0%)
missing	258 (2.6%)
Hypertension (yes)	1381 (14.0%)
missing	2 (0.0%)
Diabetes (yes)	416 (4.2%)
Year of cohort entry	
cycle1 (1994-1995)	6915 (69.9%)
cycle2 (1996-1997)	479 (4.8%)
cycle3 (1998-1999)	518 (5.2%)
cycle4 (2000-2001)	535 (5.4%)
cycle5 (2002-2003)	451 (4.6%)
cycle6 (2004-2005)	382 (3.9%)
cycle7 (2006-2007)	331 (3.4%)
cycle8 (2008-2009)	278 (2.8%)
Alcohol intake(drinks/month)	
0	6531 (66.0%)
0.1-9.9	1553 (15.7%)
10-19.9	531 (5.4%)
20-59.9	578 (5.8%)
>=60	279 (2.8%)
missing	417 (4.2%)

Abbreviations: BMI=Body mass index, IQR=interquartile range, SD=standard deviation.

*Baseline characteristics were assessed at the time of cohort entry.

[a]All frequency counts were in accordance with the release rule of Statistics Canada Research Data Center at Western University.

[b]Education was defined as highest education attainment on the cohort entry date.

[c]Income was categorized as Lowest income:(Less than $15,000 and 1 or 2 persons; Less than $20,000 and 3 or 4 persons; Less than $30,000 and 5 or more persons). Lower middle income($15,000 to $29,999 and 1 or 2 persons;$20,000 to $39,999 and 3 or 4 persons; $30,000 to $59,999 and 5 or more persons). Upper middle income($30,000 to $59,999 and 1 or 2 persons;$40,000 to $79,999 and 3 or 4

persons;$60,000 to $79,999 and 5 or more persons) Highest income($60,000 or more and 1 or 2 persons; $80,000 or more and 3 persons or more).

[d]Physical activity was categorized as being "Active", "Moderate", or "Inactive" based on the total daily EE values (kcal/kg/day) calculated for PACnDEE (PACnDEE >= 3 Active; 1.5 <= PACnDEE < 3.0 Moderate; Inactive).

Table 2. Baseline characteristics of participants across alcohol intake categories (1994-2009)

Characteristic*	Alcohol intake categories for drinker(drinks/month)[a,e]				
	0 (N=6,531)	**0.1-9.9 (n=1,553)**	**10-19.9 (n=531)**	**20-59.9 (n=578)**	**>=60 (n=279)**
Age, years, mean (SD)	55.5 (13.5)	42.7 (5.8)	49.5 (10.2)	54.5 (12.3)	58.3 (11.4)
Age groups					
40-49	2458 (37.6%)	1154 (74.3%)	289 (54.4%)	223 (38.6%)	71 (25.4%)
50-59	1124 (17.2%)	86 (5.5%)	97 (18.3%)	122 (21.1%)	60 (21.5%)
60+	2100 (32.2%)	45 (2.9%)	81 (15.3%)	163 (28.2%)	118 (42.3%)
Sex (Female)	3607 (55.2%)	599 (38.6%)	176 (33.1%)	157 (27.2%)	49 (17.6%)
Education[b]					
less than secondary education	2044 (31.3%)	168 (10.8%)	95 (17.9%)	124 (21.5%)	89 (31.9%)
secondary graduation	745 (11.4%)	178 (11.5%)	57 (10.7%)	69 (11.9%)	34 (12.2%)
some post-secondary	1272 (19.5%)	354 (22.8%)	102 (19.2%)	121 (20.9%)	41 (14.7%)
post-secondary	1621 (24.8%)	585 (37.7%)	213 (40.1%)	194 (33.6%)	85 (30.5%)
Income[c]					
lowest income	1285 (19.7%)	93 (6.0%)	51 (9.6%)	61 (10.6%)	34 (12.2%)
lower middle income	1761 (27.0%)	210 (13.5%)	93 (17.5%)	113 (19.6%)	76 (27.2%)
upper middle income	1847 (28.3%)	475 (30.6%)	192 (36.2%)	209 (36.2%)	89 (31.9%)
highest income	789 (12.1%)	507 (32.6%)	131 (24.7%)	125 (21.6%)	50 (17.9%)
Race (White)	5393 (82.6%)	1250 (80.5%)	465 (87.6%)	498 (86.2%)	246 (88.2%)
Smoking (yes)	1404 (21.5%)	454 (29.2%)	161 (30.3%)	182 (31.5%)	109 (39.1%)
BMI (kg/m^2), mean (SD)	26.2 (4.8)	26.2 (4.3)	26.2 (4.4)	25.8 (4.0)	25.9 (4.0)
Physical activity[d]					
active	808 (12.4%)	269 (17.3%)	89 (16.8%)	108 (18.7%)	48 (17.2%)
moderate	1270 (19.4%)	346 (22.3%)	106 (20.0%)	122 (21.1%)	47 (16.8%)
inactive	3604 (55.2%)	670 (43.1%)	272 (51.2%)	278 (48.1%)	154 (55.2%)
Use of vitamin (yes)	2057 (31.5%)	474 (30.5%)	187 (35.2%)	187 (32.4%)	71 (25.4%)
Hypertension (yes)	921 (14.1%)	79 (5.1%)	57 (10.7%)	83 (14.4%)	53 (19.0%)
Diabetes (yes)	286 (4.4%)	23 (1.5%)	10 (1.9%)	17 (2.9%)	17 (6.1%)

Abbreviations: BMI=Body mass index, IQR=interquartile range, SD=standard deviation.
*Baseline characteristics were assessed at the time of cohort entry.
[a]All frequency counts were in accordance with the release rule of Statistics Canada Research Data Center at Western University.
[b]Education was defined as the highest education attainment on the cohort entry date.
[c]Income was categorized as Lowest income:(Less than $15,000 and 1 or 2 persons; Less than $20,000 and 3 or 4 persons; Less than $30,000 and 5 or more persons). Lower middle income($15,000 to $29,999 and 1 or 2 persons;$20,000 to $39,999 and 3 or 4 persons; $30,000 to $59,999 and 5 or more persons). Upper middle income($30,000 to $59,999 and 1 or 2 persons;$40,000 to $79,999 and 3 or 4 persons;$60,000 to $79,999 and 5 or more persons) Highest income($60,000 or more and 1 or 2 persons; $80,000 or more and 3 persons or more).
[d]Physical activity was categorized as being "Active", "Moderate", or "Inactive" based on the total daily EE values (kcal/kg/day) calculated for PACnDEE.(PACnDEE >= 3 Active; 1.5 <= PACnDEE < 3.0 Moderate; Inactive).
[e]Frequency of diabetes across categories was not presented due to the small cell rule(<15).

Table 3. Baseline characteristics of participants aged 40 years or older from National Population Health Survey (2002-2009)

Characteristic*	Total Population (n=7,388)[a]
Age, years, mean (SD)	55.0 (13.4)
Age, years, median (IQR)	52 (21)
Age groups	
40-49	3226 (43.7%)
50-59	1713 (23.2%)
60-69	1179 (16.0%)
70+	1270 (17.2%)
Sex (Female)	4085 (55.3%)
Education[b]	
less than secondary education	1690 (22.9%)
secondary graduation	974 (13.2%)
some post-secondary	1817 (24.6%)
post-secondary	2708 (36.7%)
missing	199 (2.7%)
Income[c]	
lowest income	702 (9.5%)
lower middle income	1426 (19.3%)
upper middle income	2316 (31.4%)
highest income	2192 (29.7%)
missing	752 (10.2%)
Race	
Asian	229 (3.1%)
Black	73 (1.0%)
White	6858 (92.8%)
Other	61 (0.8%)
missing	167 (2.3%)
Smoking (yes)	1650 (22.3%)
missing	102 (1.4%)
Alcohol intake(drinks/month)	
0	4051 (54.8%)

0.1-9.9	1279 (17.3%)
10-19.9	164 (2.2%)
>=20	96 (1.3%)
missing	1789 (24.3%)
BMI (kg/m^2), mean (SD)	26.8 (4.8)
missing	284 (3.84%)
Physical activity[d]	
active	1396 (18.9%)
moderate	1949 (26.4%)
inactive	3705 (50.2%)
missing	338 (4.6%)
Use of vitamin (yes)	3059 (41.4%)
missing	258 (2.4%)
Hypertension (yes)	1590 (21.5%)
missing	1 (0.2%)
Diabetes (yes)	498 (6.7%)
Missing	1 (0.0%)
Cohort entry	
Cycle5 (2002-2003)	6180 (83.7%)
Cycle6 (2004-2005)	522 (7.1%)
Cycle7 (2006-2007)	387 (5.2%)
Cycle8 (2008-2009)	299 (4.1%)
FV intake, servings, mean (SD)	4.3 (2.1)
FV intake, servings, median (IQR)	4.0 (2.4)
FV intake groups(servings/day)	
0-2.9	1964 (26.5%)
3-5.9	3811 (51.6%)
>=6	1255 (17.0%)
missing	358 (4.9%)

Abbreviations: BMI=Body mass index, IQR=interquartile range, SD=standard deviation.

*Baseline characteristics were assessed at the time of cohort entry.

[a]All frequency counts were in accordance with the release rule of Statistics Canada Research Data Center at Western University.

[b]Education was defined as highest education attainment on the cohort entry date.
[c]Income was categorized as Lowest income:(Less than $15,000 and 1 or 2 persons; Less than $20,000 and 3 or 4 persons; Less than $30,000 and 5 or more persons). Lower middle income($15,000 to $29,999 and 1 or 2 persons;$20,000 to $39,999 and 3 or 4 persons; $30,000 to $59,999 and 5 or more persons). Upper middle income($30,000 to $59,999 and 1 or 2 persons;$40,000 to $79,999 and 3 or 4 persons;$60,000 to $79,999 and 5 or more persons) Highest income($60,000 or more and 1 or 2 persons; $80,000 or more and 3 persons or more).
[d]Physical activity was categorized as being "Active", "Moderate", or "Inactive" based on the total daily EE values (kcal/kg/day) calculated for PACnDEE.(PACnDEE >= 3 Active; 1.5 <= PACnDEE < 3.0 Moderate; Inactive).

Table 4. Baseline characteristics of participants across FV intake categories (2002-2009)

Characteristic*	FV intake categories(servings/day)[a,e]		
	0-2.9 (n=1,964)	**3-5.9 (n=3,811)**	**>=6 (n=1,255)**
Age, years, mean (SD)	51.9 (11.5)	54.6 (12.7)	54.5 (13.0)
Age groups			
40-49	954 (48.6%)	1463 (38.4%)	504 (40.2%)
50-59	427 (21.7%)	861 (22.6%)	264 (21.0%)
60-69	229 (11.7%)	580 (15.2%)	185 (14.7%)
70+	199 (10.1%)	509 (13.4%)	182 (14.5%)
Sex (Female)	790 (40.2%)	1917 (50.3%)	791 (63.0%)
Education[b]			
less than secondary education	448 (22.8%)	709 (18.6%)	205 (16.3%)
secondary graduation	296 (15.1%)	458 (12.0%)	112 (8.9%)
some post-secondary	475 (24.2%)	854 (22.4%)	316 (25.2%)
post-secondary	590 (30.0%)	1392 (36.5%)	502 (40.0%)
Income[c]			
lowest income	230 (11.7%)	298 (7.8%)	89 (7.1%)
lower middle income	410 (20.9%)	707 (18.6%)	221 (17.6%)
upper middle income	625 (31.8%)	1208 (31.7%)	414 (33.0%)
highest income	544 (27.7%)	1200 (31.5%)	411 (32.7%)
Race (White)	1721 (87.6%)	3252 (85.3%)	1100 (87.6%)
Smoking (yes)	632 (32.2%)	678 (17.8%)	166 (13.2%)
BMI (kg/m^2), mean (SD)	27.2 (5.0)	26.8 (4.7)	26.5 (4.8)
Physical activity[d]			
active	250 (12.7%)	704 (18.5%)	327 (26.1%)
moderate	422 (21.5%)	1017 (26.7%)	348 (27.7%)
inactive	1137 (57.9%)	1692 (44.4%)	460 (36.7%)
Use of vitamin (yes)	614 (31.3%)	1544 (40.5%)	554 (44.1%)
Hypertension (yes)	340 (17.3%)	724 (19.0%)	250 (19.9%)
Diabetes (yes)	103 (5.2%)	228 (6.0%)	76 (6.1%)

Abbreviations: BMI=Body mass index, IQR=interquartile range, SD=standard deviation.

*Baseline characteristics were assessed at the time of cohort entry.

[a]All frequency counts were in accordance with the release rule of Statistics Canada Research Data Center at Western University.

[b]Education was defined as the highest education attainment on the cohort entry date.

[c]Income was categorized as Lowest income:(Less than $15,000 and 1 or 2 persons; Less than $20,000 and 3 or 4 persons; Less than $30,000 and 5 or more persons). Lower middle income($15,000 to $29,999 and 1 or 2 persons;$20,000 to $39,999 and 3 or 4 persons; $30,000 to $59,999 and 5 or more persons). Upper middle income($30,000 to $59,999 and 1 or 2 persons;$40,000 to $79,999 and 3 or 4 persons;$60,000 to $79,999 and 5 or more persons) Highest income($60,000 or more and 1 or 2 persons; $80,000 or more and 3 persons or more).

[d]Physical activity was categorized as being "Active", "Moderate", or "Inactive" based on the total daily EE values (kcal/kg/day) calculated for PACnDEE.(PACnDEE >= 3 Active; 1.5 <= PACnDEE < 3.0 Moderate; Inactive).

[e]Frequency of diabetes across categories was not presented due to the small cell rule(<15).

Table 5. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for NPHS respondents (1994-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.99 to 1.04)	0.38	1.02 (0.99 to 1.04)	0.15
Age (years)			1.08 (1.08 to 1.09)	<.0001¶
Sex (female)			1.39 (1.25 to 1.55)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.89 (0.75 to 1.05)	0.17
some post-secondary graduation			1.18 (1.04 to 1.35)	0.01¶
post-secondary graduation			0.94 (0.82 to 1.08)	0.36
Income				
lowest income			Reference	
lower middle income			1.03 (0.91 to 1.18)	0.64
upper middle income			1.09 (0.94 to 1.26)	0.25
highest income			1.09 (0.89 to 1.33)	0.40
Race				
White			Reference	
Asian			1.20 (0.87 to 1.67)	0.27
Black			0.79 (0.39 to 1.58)	0.50
Other			1.33 (0.59 to 2.98)	0.49
Smoking (yes)			1.11 (0.98 to 1.26)	0.11
BMI (kg/m^2)			1.01 (1.00 to 1.02)	0.07
Physical -activity				
active			Reference	
moderate			1.14 (0.97 to 1.34)	0.13
inactive			1.03 (0.89 to 1.19)	0.69
Use of vitamin (yes)			1.10 (1.00 to 1.22)	0.06
Hypertension (yes)			1.11 (0.99 to 1.25)	0.09
Diabetes (yes)			1.78 (1.48 to 2.14)	<.0001¶

Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.45 (0.26 to 0.78)	<.01¶
cycle3 (1998-1999)	0.46 (0.27 to 0.79)	<.01¶
cycle4 (2000-2001)	0.32 (0.14 to 0.71)	<.01¶
cycle5 (2002-2003)	0.87 (0.46 to 1.64)	0.67
cycle6 (2004-2005)	0.14 (0.02 to 1.00)	0.05
cycle7 (2006-2007)	0.51 (0.13 to 2.06)	0.35
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 6. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts including quadratic terms for NPHS respondents (1994-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.99 to 1.04)	0.38	1.00 (0.95 to 1.04)	0.83
Alcohol intake^2 (10 drinks/month)			1.001 (0.999 to 1.004)	0.34
Age (years)			1.31 (1.25 to 1.37)	<.0001¶
Age^2 (years)			0.999 (0.998 to 0.999)	<.0001¶
Sex (female)			1.39 (1.24 to 1.55)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.90 (0.76 to 1.06)	0.19
some post-secondary graduation			1.21 (1.06 to 1.38)	<.01¶
post-secondary graduation			0.97 (0.85 to 1.12)	0.70
Income				
lowest income			Reference	
lower middle income			0.99 (0.87 to 1.13)	0.91
upper middle income			1.06 (0.92 to 1.22)	0.43
highest income			1.07 (0.88 to 1.30)	0.51
Race				
White			Reference	
Asian			1.22 (0.88 to 1.68)	0.24
Black			0.83 (0.41 to 1.67)	0.60
Other			1.36 (0.61 to 3.05)	0.46
Smoking (yes)			1.08 (0.95 to 1.23)	0.25
BMI (kg/m^2)			0.96 (0.89 to 1.04)	0.33
BMI^2 (kg/m^2)			1.001 (0.999 to 1.002)	0.27
Physical activity				
active			Reference	
moderate			1.13 (0.96 to 1.32)	0.16
inactive			1.06 (0.92 to 1.22)	0.43

Use of vitamin (yes)	1.08 (0.97 to 1.19)	0.15
Hypertension (yes)	1.09 (0.97 to 1.23)	0.14
Diabetes (yes)	1.75 (1.46 to 2.11)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.82 (0.46 to 1.47)	0.51
cycle3 (1998-1999)	0.84 (0.48 to 1.49)	0.54
cycle4 (2000-2001)	0.57 (0.25 to 1.31)	0.19
cycle5 (2002-2003)	1.60 (0.83 to 3.07)	0.16
cycle6 (2004-2005)	0.25 (0.04 to 1.81)	0.17
cycle7 (2006-2007)	0.92 (0.23 to 3.75)	0.91
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 7. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts with time-varying covariates for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)#	1.05 (1.00 to 1.09)	0.05	1.03 (0.98 to 1.09)	0.25
Age (years)			1.09 (1.08 to 1.11)	<.0001¶
Sex (female)			0.95 (0.70 to 1.28)	0.71
Education				
less than secondary graduation			Reference	
secondary graduation			1.04 (0.66 to 1.65)	0.86
some post-secondary graduation			1.26 (0.88 to 1.81)	0.22
post-secondary graduation			0.97 (0.64 to 1.47)	0.88
Income				
lowest income			Reference	
lower middle income			0.93 (0.67 to 1.28)	0.66
upper middle income			0.65 (0.43 to 0.99)	0.04¶
highest income			0.51 (0.25 to 1.04)	0.06
Race				
White			Reference	
Asian			0.75 (0.24 to 2.35)	0.62
Black			N/A	N/A
Other			N/A	N/A
Smoking (yes)			1.23 (0.85 to 1.78)	0.27
BMI (kg/m^2) #			1.02 (0.99 to 1.05)	0.21
Physical- activity				
active			Reference	
moderate			1.19 (0.72 to 1.97)	0.50
inactive			1.25 (0.81 to 1.93)	0.31
Use of vitamin (yes)			1.18 (0.89 to 1.57)	0.24
Hypertension (yes)			1.12 (0.82 to 1.53)	0.49

Diabetes (yes)	1.84 (1.20 to 2.81)	0.01¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.56 (0.08 to 4.14)	0.57
cycle3 (1998-1999)	1.42 (0.28 to 7.02)	0.45
cycle4 (2000-2001)	N/A	N/A
cycle5 (2002-2003)	1.60 (0.38 to 6.78)	0.53
cycle6 (2004-2005)	N/A	N/A
cycle7 (2006-2007)	1.12 (0.15 to 8.31)	0.91
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

Included as a time-varying covariate

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 8. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for NPHS respondents (2002-2011) with multiple imputations (30 imputations)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.99 to 1.04)	0.36	1.02 (1.00 to 1.04)	0.06
Age (years)			1.08 (1.08 to 1.09)	<.0001¶
Sex (female)			1.40 (1.27 to 1.54)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.88 (0.75 to 1.02)	0.08
some post-secondary graduation			1.11 (0.98 to 1.25)	0.10
post-secondary graduation			0.89 (0.78 to 1.01)	0.07
Income				
lowest income			Reference	
lower middle income			1.03 (0.91 to 1.16)	0.67
upper middle income			1.07 (0.94 to 1.23)	0.30
highest income			1.10 (0.91 to 1.32)	0.33
Race				
White			Reference	
Asian			1.20 (0.90 to 1.60)	0.22
Black			0.71 (0.36 to 1.40)	0.32
Other			1.13 (0.57 to 2.25)	0.73
Smoking (yes)			1.11 (0.99 to 1.25)	0.07
BMI (kg/m^2)			1.01 (1.00 to 1.02)	0.02¶
Physical activity				
active			Reference	
moderate			1.14 (0.97 to 1.33)	0.10
inactive			1.01 (0.88 to 1.15)	0.91
Use of vitamin (yes)			1.11 (1.01 to 1.22)	0.03¶
Hypertension (yes)			1.13 (1.02 to 1.26)	0.02¶
Diabetes (yes)			1.63 (1.39 to 1.93)	<.0001¶

Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.43 (0.25 to 0.74)	<.01¶
cycle3 (1998-1999)	0.44 (0.25 to 0.77)	<.01¶
cycle4 (2000-2001)	0.26 (0.12 to 0.58)	<.01¶
cycle5 (2002-2003)	0.72 (0.39 to 1.31)	0.28
cycle6 (2004-2005)	0.21 (0.05 to 0.86)	0.03¶
cycle7 (2006-2007)	0.37 (0.09 to 1.48)	0.16
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 9. Crude and adjusted Hazard Ratio for the association between alcohol intake -and cataracts with time-varying covariates for NPHS respondents (2002-2011) with multiple imputations (30 imputations)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month) #	1.05 (1.00 to 1.10)	0.04¶	0.99 (0.90 to 1.08)	0.83
Age (years)			1.25 (1.12 to 1.39)	<.0001¶
Sex (female)			0.91 (0.71 to 1.17)	0.45
Education				
less than secondary graduation			Reference	
secondary graduation			1.04 (0.71 to 1.53)	0.82
some post-secondary graduation			1.12 (0.82 to 1.53)	0.49
post-secondary graduation			0.88 (0.61 to 1.26)	0.48
Income				
lowest income			Reference	
lower middle income			0.84 (0.64 to 1.10)	0.20
upper middle income			0.67 (0.47 to 0.96)	0.03¶
highest income			0.71 (0.40 to 1.25)	0.24
Race				
White			Reference	
Asian			0.75 (0.30 to 1.91)	0.54
Black			N/A	N/A
Other			1.58 (0.51 to 4.88)	0.43
Smoking (yes)			1.20 (0.88 to 1.64)	0.25
BMI (kg/m^2) #			1.00 (0.86 to 1.17)	0.99
Physical activity				
active			Reference	
moderate			1.35 (0.85 to 2.14)	0.20
inactive			1.35 (0.90 to 2.00)	0.14
Use of vitamin (yes)			1.06 (0.84 to 1.36)	0.62
Hypertension (yes)			1.04 (0.80 to 1.35)	0.77

Diabetes (yes)	1.51 (1.06 to 2.15)	0.02¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.89 (0.11 to 7.02)	0.91
cycle3 (1998-1999)	1.61 (0.35 to 7.48)	0.54
cycle4 (2000-2001)	N/A	N/A
cycle5 (2002-2003)	2.04 (0.44 to 9.42)	0.36
cycle6 (2004-2005)	N/A	N/A
cycle7 (2006-2007)	1.36 (0.17 to 10.75)	0.77
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

Included as a time-varying covariate

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 10. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts with excluding outliers of alcohol intake for NPHS respondents (1994-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.99 to 1.04)	0.38	1.00 (0.95 to 1.04)	0.83
Alcohol intake^2 (10 drinks/month)			1.001 (0.999 to 1.004)	0.34
Age (years)			1.31 (1.25 to 1.37)	<.0001¶
Age^2 (years)			0.999 (0.998 to 0.999)	<.0001¶
Sex (female)			1.39 (1.24 to 1.55)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.90 (0.76 to 1.06)	0.19
some post-secondary graduation			1.21 (1.06 to 1.38)	<.01¶
post-secondary graduation			0.97 (0.85 to 1.12)	0.70
Income				
lowest income			Reference	
lower middle income			0.99 (0.87 to 1.13)	0.91
upper middle income			1.06 (0.92 to 1.22)	0.43
highest income			1.07 (0.88 to 1.30)	0.51
Race				
White			Reference	
Asian			1.22 (0.88 to 1.68)	0.24
Black			0.83 (0.41 to 1.67)	0.60
Other			1.36 (0.61 to 3.05)	0.46
Smoking (yes)			1.08 (0.95 to 1.23)	0.25
BMI (kg/m^2)			0.96 (0.89 to 1.04)	0.33
BMI^2 (kg/m^2)			1.001 (0.999 to 1.002)	0.27
Physical activity				
active			Reference	
moderate			1.13 (0.96 to 1.32)	0.16
inactive			1.06 (0.92 to 1.22)	0.43

Use of vitamin (yes)	1.08 (0.97 to 1.19)	0.15
Hypertension (yes)	1.09 (0.97 to 1.23)	0.14
Diabetes (yes)	1.75 (1.46 to 2.11)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.82 (0.46 to 1.47)	0.51
cycle3 (1998-1999)	0.84 (0.48 to 1.49)	0.54
cycle4 (2000-2001)	0.57 (0.25 to 1.31)	0.19
cycle5 (2002-2003)	1.60 (0.83 to 3.07)	0.16
cycle6 (2004-2005)	0.25 (0.04 to 1.81)	0.17
cycle7 (2006-2007)	0.92 (0.23 to 3.75)	0.91
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 11. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for white NPHS respondents (1994-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.98 to 1.04)	0.47	1.00 (0.95 to 1.05)	0.95
Alcohol intake^2 (10 drinks/month)			1.001 (0.998 to 1.004)	0.39
Age (years)			1.32 (1.26 to 1.38)	<.0001¶
Age^2 (years)			0.998 (0.998 to 0.999)	<.0001¶
Sex (female)			1.42 (1.27 to 1.59)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.91 (0.77 to 1.08)	0.29
some post-secondary graduation			1.22 (1.07 to 1.40)	<.01¶
post-secondary graduation			0.98 (0.85 to 1.13)	0.75
Income				
lowest income			Reference	
lower middle income			0.96 (0.85 to 1.10)	0.58
upper middle income			1.03 (0.89 to 1.19)	0.71
highest income			1.01 (0.82 to 1.24)	0.94
Smoking (yes)			1.07 (0.94 to 1.22)	0.31
BMI (kg/m^2)			0.96 (0.89 to 1.04)	0.31
BMI^2 (kg/m^2)			1.001 (0.999 to 1.002)	0.26
Physical activity				
active			Reference	
moderate			1.14 (0.97 to 1.35)	0.12
inactive			1.05 (0.91 to 1.22)	0.48
Use of vitamin (yes)			1.08 (0.97 to 1.20)	0.15
Hypertension (yes)			1.08 (0.95 to 1.21)	0.24

Diabetes (yes)	1.75 (1.45 to 2.12)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.75 (0.40 to 1.40)	0.36
cycle3 (1998-1999)	0.77 (0.41 to 1.43)	0.37
cycle4 (2000-2001)	0.62 (0.27 to 1.43)	0.26
cycle5 (2002-2003)	1.62 (0.82 to 3.22)	0.17
cycle6 (2004-2005)	0.28 (0.04 to 2.01)	0.21
cycle7 (2006-2007)	1.00 (0.25 to 4.07)	1.00
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 12. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.06 (1.04 to 1.09)	<.0001¶	1.00 (0.96 to 1.05)	1.00
Alcohol intake (10 drinks/month)	0.53 (0.39 to 0.73)	<.0001¶	0.95 (0.76 to 1.17)	0.60
Age (years)			1.08 (1.07 to 1.09)	<.0001¶
Sex (female)			1.59 (1.32 to 1.92)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.91 (0.68 to 1.22)	0.27
some post-secondary graduation			1.14 (0.90 to 1.45)	0.79
post-secondary graduation			0.97 (0.76 to 1.23)	0.25
Income				
lowest income			Reference	
lower middle income			1.00 (0.74 to 1.33)	0.78
upper middle income			1.04 (0.78 to 1.40)	0.64
highest income			0.92 (0.66 to 1.30)	0.64
Race				
White			Reference	
Asian			1.00 (0.50 to 2.03)	0.99
Black			1.04 (0.26 to 4.21)	0.95
Other			1.16 (0.37 to 3.65)	0.80
Smoking (yes)			1.15 (0.91 to 1.46)	0.25
BMI (kg/m^2)			1.02 (1.00 to 1.04)	0.29
Physical activity				
active			Reference	
moderate			1.11 (0.86 to 1.42)	0.42

inactive	0.92 (0.72 to 1.16)	0.47
Use of vitamin (yes)	1.10 (0.92 to 1.31)	0.34
Hypertension (yes)	1.10 (0.91 to 1.33)	0.34
Diabetes (yes)	1.18 (0.88 to 1.59)	0.26
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	0.34 (0.08 to 1.36)	0.75
cycle7 (2006-2007)	0.83 (0.26 to 2.63)	0.95
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 13. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts including quadratic terms for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.06 (1.04 to 1.09)	<.0001¶	1.04 (0.91 to 1.19)	0.52
Fruit and vegetable consumption^2 (servings/day)			1.00 (0.99 to 1.01)	0.47
Alcohol intake (10 drinks/month)			0.92 (0.64 to 1.33)	0.65
Alcohol intake^2 (10 drinks/month)			1.01 (0.96 to 1.06)	0.78
Age (years)			1.49 (1.37 to 1.62)	<.0001¶
Age^2 (years)			0.998 (0.997 to 0.998)	<.0001¶
Sex (female)			1.60 (1.32 to 1.93)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.96 (0.71 to 1.29)	0.77
some post-secondary graduation			1.20 (0.95 to 1.52)	0.14
post-secondary graduation			1.04 (0.82 to 1.33)	0.73
Income				
lowest income			Reference	
lower middle income			0.96 (0.72 to 1.30)	0.81
upper middle income			1.00 (0.74 to 1.34)	0.99
highest income			0.88 (0.62 to 1.23)	0.44
Race				
White			Reference	
Asian			0.92 (0.45 to 1.86)	0.82
Black			1.08 (0.27 to 4.35)	0.92
Other			1.84 (0.59 to 5.80)	0.30
Smoking (yes)			1.13 (0.89 to 1.43)	0.31
BMI (kg/m^2)			1.04 (0.89 to 1.20)	0.65
BMI^2 (kg/m2)			1.000 (0.997 to 1.002)	0.74

Physical activity		
active	Reference	
moderate	1.14 (0.89 to 1.46)	0.31
inactive	0.97 (0.76 to 1.22)	0.77
Use of vitamin (yes)	1.07 (0.90 to 1.28)	0.42
Hypertension (yes)	1.02 (0.85 to 1.24)	0.81
Diabetes (yes)	1.15 (0.86 to 1.54)	0.34
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.04 (0.25 to 4.40)	0.96
cycle7 (2006-2007)	2.64 (0.79 to 8.76)	0.11
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 14. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts with time-varying covariates for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day) #	1.07 (1.01 to 1.12)	0.01¶	1.03 (0.95 to 1.12)	0.42
Alcohol intake (10 drinks/month) #			0.99 (0.66 to 1.48)	0.97
Age (years)			1.09 (1.08 to 1.11)	<.0001¶
Sex (female)			1.50 (1.05 to 2.15)	0.03¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.84 (0.50 to 1.42)	0.52
some post-secondary graduation			0.75 (0.48 to 1.19)	0.22
post-secondary graduation			0.73 (0.46 to 1.14)	0.16
Income				
lowest income			Reference	
lower middle income			1.02 (0.60 to 1.72)	0.95
upper middle income			1.22 (0.71 to 2.08)	0.47
highest income			0.89 (0.46 to 1.73)	0.73
Race				
White			Reference	
Asian			2.47 (0.90 to 6.83)	0.08
Black			N/A	N/A
Other			1.16 (0.16 to 8.66)	0.89
Smoking (yes)			1.39 (0.89 to 2.19)	0.15
BMI (kg/m^2) #			1.04 (1.00 to 1.08)	0.03¶
Physical activity				
active			Reference	
moderate			1.03 (0.64 to 1.66)	0.91
inactive			0.81 (0.51 to 1.27)	0.35
Use of vitamin (yes)			1.07 (0.77 to 1.50)	0.69

Hypertension (yes)	1.15 (0.81 to 1.64)	0.44
Diabetes (yes)	1.21 (0.72 to 2.03)	0.47
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	0.78 (0.10 to 5.76)	0.80
cycle7 (2006-2007)	1.60 (0.38 to 6.81)	0.52
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

Included as a time-varying covariate

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 15. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts for NPHS respondents (2002-2011) with multiple imputations (30 imputations)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.06 (1.03 to 1.09)	<.0001¶	1.01 (0.98 to 1.04)	0.51
Alcohol intake (10 drinks/month)			0.94 (0.77 to 1.14)	0.53
Age (years)			1.07 (1.07 to 1.08)	<.0001¶
Sex (female)			1.32 (1.15 to 1.52)	<.01¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.92 (0.75 to 1.15)	0.47
some post-secondary graduation			1.17 (0.99 to 1.40)	0.07
post-secondary graduation			0.90 (0.75 to 1.07)	0.23
Income				
lowest income			Reference	
lower middle income			1.09 (0.89 to 1.33)	0.40
upper middle income			1.04 (0.84 to 1.30)	0.70
highest income			0.93 (0.72 to 1.21)	0.60
Race				
White			Reference	
Asian			1.07 (0.71 to 1.62)	0.74
Black			0.63 (0.24 to 1.68)	0.36
Other			1.75 (0.86 to 3.56)	0.12
Smoking (yes)			1.10 (0.92 to 1.32)	0.30
BMI (kg/m^2)			1.01 (1.00 to 1.03)	0.06
Physical activity				
active			Reference	
moderate			1.02 (0.83 to 1.25)	0.85
inactive			0.95 (0.79 to 1.15)	0.60

Use of vitamin (yes)	1.17 (1.02 to 1.33)	0.02¶
Hypertension (yes)	1.18 (1.04 to 1.36)	0.01¶
Diabetes (yes)	1.40 (1.16 to 1.69)	<.01¶
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	0.73 (0.48 to 1.11)	0.14
cycle7 (2006-2007)	0.69 (0.33 to 1.47)	0.34
cycle8 (2008-2009)	0.33 (0.05 to 2.36)	0.27

Abbreviations: CI=Confidence interval, BMI=Body mass index

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 16. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts with time-varying covariates for NPHS respondents (2002-2011) with multiple imputations (30 imputations)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day) #	1.06 (1.01 to 1.11)	0.02¶	0.99 (0.89 to 1.11)	0.92
Alcohol intake (10 drinks/month) #			0.93 (0.55 to 1.57)	0.78
Alcohol intake^2 (10 drinks/month)			1.01 (0.92 to 1.10)	0.86
Age (years)			1.39 (1.25 to 1.53)	<.0001¶
Sex (female)			1.33 (1.02 to 1.73)	0.04¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.94 (0.63 to 1.40)	0.76
some post-secondary graduation			1.12 (0.82 to 1.52)	0.48
post-secondary graduation			0.86 (0.61 to 1.22)	0.40
Income				
lowest income			Reference	
lower middle income			1.12 (0.79 to 1.58)	0.53
upper middle income			1.11 (0.77 to 1.61)	0.58
highest income			0.81 (0.49 to 1.35)	0.42
Race				
White			Reference	
Asian			1.08 (0.47 to 2.46)	0.86
Black			0.61 (0.09 to 4.30)	0.62
Other			1.64 (0.41 to 6.64)	0.49
Smoking (yes)			1.12 (0.79 to 1.60)	0.52
BMI (kg/m^2) #			1.15 (0.93 to 1.41)	0.19
Physical activity				
active			Reference	
moderate			0.95 (0.64 to 1.41)	0.79

inactive	0.90 (0.63 to 1.28)	0.57
Use of vitamin (yes)	1.30 (1.02 to 1.65)	0.03¶
Hypertension (yes)	1.18 (0.92 to 1.50)	0.19
Diabetes (yes)	1.17 (0.84 to 1.63)	0.35
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.12 (0.57 to 2.20)	0.39
cycle7 (2006-2007)	1.56 (0.57 to 4.31)	0.97
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

Included as a time-varying covariate

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 17. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts with excluding outliers of FV intake for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.07 (1.04 to 1.11)	<.0001¶	1.09 (0.88 to 1.34)	0.43
Alcohol intake (10 drinks/month)			0.92 (0.64 to 1.33)	0.67
Alcohol intake^2 (10 drinks/month)			1.01 (0.96 to 1.06)	0.79
Age (years)			1.49 (1.37 to 1.62)	<.0001¶
Age^2 (years)			0.998 (0.997 to 0.998)	<.0001¶
Sex (female)			1.61 (1.33 to 1.95)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.95 (0.70 to 1.27)	0.71
some post-secondary graduation			1.20 (0.95 to 1.53)	0.12
post-secondary graduation			1.05 (0.82 to 1.33)	0.72
Income				
lowest income			Reference	
lower middle income			0.96 (0.72 to 1.30)	0.80
upper middle income			1.00 (0.74 to 1.34)	0.97
highest income			0.88 (0.62 to 1.23)	0.45
Race				
White			Reference	
Asian			0.92 (0.46 to 1.87)	0.82
Black			1.09 (0.27 to 4.39)	0.91
Other			1.82 (0.58 to 5.74)	0.30
Smoking (yes)			1.14 (0.90 to 1.45)	0.27
BMI (kg/m^2)			1.04 (0.90 to 1.21)	0.60
BMI^2 (kg/m^2)			1.000 (0.997 to 1.002)	0.71
Physical activity				
active			Reference	

moderate	1.13 (0.89 to 1.45)	0.32
inactive	0.96 (0.76 to 1.22)	0.76
Use of vitamin (yes)	1.06 (0.89 to 1.27)	0.49
Hypertension (yes)	1.02 (0.84 to 1.23)	0.86
Diabetes (yes)	1.16 (0.87 to 1.56)	0.32
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.06 (0.25 to 4.48)	0.94
cycle7 (2006-2007)	2.73 (0.82 to 9.07)	0.10
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 18. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts for white NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.06 (1.03 to 1.09)	<.0001¶	1.05 (0.92 to 1.20)	0.49
Alcohol intake (10 drinks/month)			0.93 (0.64 to 1.35)	0.71
Alcohol intake^2 (10 drinks/month)			1.01 (0.96 to 1.06)	0.81
Age (years)			1.51 (1.38 to 1.65)	<.0001¶
Age^2 (years)			0.997 (0.997 to 0.998)	<.0001¶
Sex (female)			1.65 (1.36 to 2.01)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.98 (0.73 to 1.32)	0.90
some post-secondary graduation			1.19 (0.94 to 1.51)	0.15
post-secondary graduation			1.03 (0.81 to 1.32)	0.80
Income				
lowest income			Reference	
lower middle income			0.95 (0.70 to 1.27)	0.72
upper middle income			0.97 (0.72 to 1.31)	0.86
highest income			0.87 (0.62 to 1.23)	0.44
Smoking (yes)			1.13 (0.89 to 1.43)	0.32
BMI (kg/m^2)			1.04 (0.89 to 1.20)	0.64
BMI^2 (kg/m^2)			1.000 (0.997 to 1.002)	0.74
Physical activity				
active			Reference	
moderate			1.10 (0.86 to 1.41)	0.45
inactive			0.93 (0.74 to 1.18)	0.56
Use of vitamin (yes)			1.08 (0.90 to 1.28)	0.42
Hypertension (yes)			1.03 (0.85 to 1.25)	0.75
Diabetes (yes)			1.18 (0.88 to 1.58)	0.26

Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.20 (0.28 to 5.08)	0.81
cycle7 (2006-2007)	3.00 (0.90 to 10.04)	0.07
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 19. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	0.53 (0.39 to 0.73)	<.0001¶	0.93 (0.64 to 1.34)	0.69
Alcohol intake^2 (10 drinks/month)			1.01 (0.96 to 1.06)	0.82
Age (years)			1.49 (1.37 to 1.62)	<.0001¶
Age^2 (years)			0.998 (0.997 to 0.998)	<.0001¶
Sex (female)			1.59 (1.33 to 1.92)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.98 (0.73 to 1.31)	0.87
some post-secondary graduation			1.22 (0.97 to 1.55)	0.09
post-secondary graduation			1.06 (0.84 to 1.35)	0.63
Income				
lowest income			Reference	
lower middle income			0.96 (0.72 to 1.28)	0.78
upper middle income			0.99 (0.74 to 1.32)	0.92
highest income			0.86 (0.62 to 1.21)	0.39
Race				
White			Reference	
Asian			1.01 (0.52 to 1.97)	0.97
Black			1.00 (0.25 to 4.03)	1.00
Other			1.87 (0.59 to 5.86)	0.29
Smoking (yes)			1.13 (0.89 to 1.42)	0.32
BMI (kg/m^2)			1.04 (0.90 to 1.19)	0.64
BMI^2 (kg/m^2)			1.000 (0.997 to 1.002)	0.74
Physical activity				
active			Reference	
moderate			1.15 (0.90 to 1.47)	0.26
inactive			0.98 (0.78 to 1.24)	0.87

Use of vitamin (yes)	1.08 (0.91 to 1.28)	0.40
Hypertension (yes)	1.04 (0.86 to 1.25)	0.72
Diabetes (yes)	1.16 (0.87 to 1.55)	0.31
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.04 (0.25 to 4.39)	0.96
cycle7 (2006-2007)	2.60 (0.78 to 8.63)	0.12
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Table 20. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts with time-varying covariates for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month) #	0.38 (0.18 to 0.78)	<.01¶	0.99 (0.66 to 1.48)	0.97
Age (years)			1.09 (1.08 to 1.11)	<.0001¶
Sex (female)			1.54 (1.08 to 2.19)	0.02¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.86 (0.51 to 1.44)	0.56
some post-secondary graduation			0.79 (0.50 to 1.24)	0.30
post-secondary graduation			0.74 (0.47 to 1.15)	0.18
Income				
lowest income			Reference	
lower middle income			1.11 (0.66 to 1.87)	0.69
upper middle income			1.31 (0.77 to 2.24)	0.32
highest income			0.96 (0.49 to 1.87)	0.91
Race				
White			Reference	
Asian			2.32 (0.84 to 6.38)	0.10
Black			N/A	N/A
Other			1.16 (0.16 to 8.67)	0.88
Smoking (yes)			1.37 (0.87 to 2.14)	0.17
BMI (kg/m^2) #			1.04 (1.00 to 1.08)	0.04¶
Physical activity				
active			Reference	
moderate			1.00 (0.62 to 1.62)	0.99
inactive			0.79 (0.51 to 1.24)	0.30
Use of vitamin (yes)			1.08 (0.78 to 1.51)	0.64
Hypertension (yes)			1.18 (0.83 to 1.67)	0.36

Diabetes (yes)	1.21 (0.72 to 2.02)	0.47
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	0.79 (0.11 to 5.84)	0.81
cycle7 (2006-2007)	1.62 (0.38 to 6.87)	0.51
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

Included as a time-varying covariate

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Chapter 6 Discussion

In this chapter, a summary of key findings of this study are provided in Section 6.1. The strengths and limitations of the study are discussed in Section 6.2 and Section 6.3. A conclusion for the results is summarized in Section 6.4.

6.1 Thesis Summary

In this study, we examined the effect of alcohol intake on the risk of cataracts using data from NPHS, a large Canadian population-based longitudinal study. We also estimated the association of fruit and vegetable consumption with cataracts. We used cycle 1 to cycle 9 of NPHS to address the association between alcohol use and cataracts. Due to the data availability, we only utilised data from cycle 5 onwards to investigate our second study objective. A comprehensive literature review was conducted, which served as the basis for establishing the study's causal framework, as represented by the corresponding directed acyclic graph (DAG). We did not find a statistically significant association between either exposure of interest and cataracts using the Cox proportional hazards model. The estimates remained robust when we performed the multiple imputations, removed outliers, treated death as a competing risk, or restricted our study population to white people. No statistically significant association was found when we performed subgroup analysis, and there was no evidence of effect modification by age or sex.

6.1.1 The Association between Alcohol Use and Cataracts

Our first objective aimed to assess if alcohol use was associated with the risk of cataracts. We hypothesized that alcohol intake increases the risk of cataracts. However, the Cox proportional models showed that cumulative alcohol use was not statistically significantly associated with cataracts. The estimation remained robust when we conducted multiple imputations using fully conditional method (FCS).

The Fine and Gray's sub-distribution hazard model suggested there was no statistically significant association between cumulative alcohol use and cataracts in the presence of competing risks. The reason behind this might be due to low death rate in our cohort (1%). We could not report the number of death due to privacy reasons of RDC center. When four subgroups of the study population were examined separately, none showed evidence of the

higher risk of cataracts with the increased use of alcohol. We also did not find evidence of effect modification of association by sex or age.

Our findings are consistent with ten previously published studies (5,7,104–109,111,114). Our study measured alcohol use as a continuous variable and calculated a cumulative amount up to the date of cohort entry. Only two published observational cohort studies measured alcohol use as a continuous variable (107,150). The study by Lindblad et al. suggested that a 13g daily intake of alcohol was associated with a 7% increased risk of cataract extraction (RR:1.07, 95% CI: 1.02-1.21). However, this study excluded missing data from the analysis, which might have distorted estimation. Furthermore, cataract extraction was used as a proxy measurement of cataracts. It could underestimate the occurrence of cataracts since only cataract patients with severe visual impairment are likely to seek cataract surgery, resulting misclassification of cataracts. Additionally, this study only measured alcohol at baseline and did not include any time-varying covariates (103).

Another study by Chasan-Taber et al. had a similar finding as our study, suggesting that the association between alcohol use and cataracts was not statistically significant. This study used a similar alcohol measurement with our study, computing cumulative alcohol intake from all questionnaires up to the start of each follow-up. Alcohol use was categorized into $\leqslant$4.9, 5.0-14.9, 15.0-24.9 and $\geqslant$25g/day, and never drinking was the reference group. Cataract extraction acted as a surrogate for cataracts in this study. Notably, cumulative alcohol use in this study included alcohol intake at the time of follow-up, which means exposure and outcome could be measured at the same time, raising a concern on causal inference due to temporality (107).

However, some contradictory findings were also reported in previous studies. The study by McCarty et al. found a statistically significant association of alcohol use with cataracts (112). Both Cumming et al. and Morris et al. found alcohol use was statistically significantly associated with a higher risk of nuclear cataracts, but not for other types of cataracts (110,113). It is necessary to note that all of these three studies are cross-sectional. Thus, causal inference based on these results is questionable as temporality cannot be examined for cross-sectional design. In addition, McCarty et al. did not control confounders at all, and Cumming et al. only adjusted for age and vitamin use in their analysis (112,113).

Both case-control studies showed a statistically significant association between alcohol consumption and cataracts (115,116). However, recall bias in case-control studies could have biased the estimation. For example, subjects with cataracts could have been more likely to report alcohol use in the reporting period.

Our results were very similar with the findings of a recent systematic review and meta-analysis. The systematic review and meta-analysis by Wang et al. of 119,706 participants from seven prospective cohort studies suggested that alcohol intake was not statistically significantly associated with an increased risk of cataracts (151).

However, a systematic review and meta-analysis by Gong et al., including five case-control studies and five cohort studies, showed heavy alcohol consumption was statistically significantly associated with cataracts (pooled relative risk [RR]: 1.26, 95% CI: 1.06 to 1.50), while the association between moderate consumption and cataracts was not statistically significant (pooled RR: 0.88, 95% CI: 0.74 to 1.05) (152).

Misclassification of cataracts could occur in our study since cataracts were identified based on a self-reported professional-diagnosed cataract. A validation study that compared self-reported eye diseases with medical records. Cohen's kappa (κ) was used to examine agreement. This study showed agreement between self-reported cataracts and medical records was not substantial (κ=0.18). Self-report tended to underestimate eye diseases (153). Misclassification in our study is likely to be non-differential, the outcome assessment occurred after the exposure examination. Thus, it could have produced a bias towards the null, which might partially explain the results we produced (154).

Another explanation for our results might be due to residual confounding (155). For example, we were not able to adjust for UV exposure, which has been recognized as a risk factor for cataracts. However, it is usually hard to be measured accurately and is rarely considered in eye disease studies. Another explanatory variable that was not available in NPHS is the health care encounter. Failure to include these two variables and other unknown confounders could have produced biased results in our study. Residual confounding may aslo arise from inappropriate categorization of continous variables due to the data availability in NPHS. For example, income

was devided into four categories, and smoking was grouped as smokers or non-smokers. Confounding could still exist after adjustment for these measured confounders.

6.1.2 The Association between FV Intake and Cataracts

The purpose of our second objective was to examine if FV intake has a protective effect on cataracts. No statistically significant association was found between FV intake and cataracts in all of our regression models. The findings were consistent when we performed several planned sensitivity analyses, including multiple imputations, excluding outliers and limiting the study population to white participants only.

The findings from our study were consistent with the results of a study by Cumming et al., which examined the association of vegetables with cataracts (130). The vegetables in that study included broccoli, brussell sprouts, cabbage, carrots, cauliflower, green beans, peas and spinach. None of them was statistically significantly associated with cataracts. Fruit intake was not investigated in that study. All vegetable intake was measured as times and categorized into five groups (<1 time/month, 1-3 times/month, 1 time/week, 2-4 times/week, >4 times/week). However, this study had a cross-sectional design and cannot establish causality of the exposure and outcome of interest.

Contrary to our findings, the study by Adachi et al. found that total vegetable consumption increased the risk of cataracts in females (OR: 1.28, 95% CI: 1.06-1.53), although it was not statistically significantly associated with cataracts in males (OR: 0.77, 95% CI: 0.59-1.01). No evidence was found for fruit intake (120). It is important to note that the OR for this study was for the highest quartile of daily intake compared with the lowest quartile. Categorization of continuous vegetable intake in this study was problematic, as multiple comparisons are likely to increase the chance of ‘false positive’ results (131).

Another study by Christen et al. also suggested that FV consumption was associated with a decreased risk of cataracts (RR for highest quintile vs lowest quintile=0.83; CI 95%: 0.79-0.99) (122). A study by Chasan-Taber et al. indicated that spinach and other green vegetables rich in carotenoids reduced the risk of cataracts, but not for other vegetables and fruit (123). Similarly, both studies categorized FV intake into quintiles, which might have produced distorted estimation, because of the implausible assumption of homogeneous risk within categories (135).

FV consumption has also been consistently indicated to have a protective effect on the cataracts by several other observational studies (115,124,125,127–129).

Oxidative stress has been suggested to play a vital role in lens opacification (69). The hypothesis that FV consumption could play a protective role in cataract development was supported by the fact that some fruit and vegetables are abundant in antioxidants, such as vitamin C, vitamin E, β-carotene, carotenoid, lutein or zeaxanthin etc. (156,157). However, some randomized controlled trials (RCTs) did not support that antioxidant vitamin supplements have a protective effect against cataracts. A recent systematic review and analysis by Mathew et al., including 117,272 participants from nine RCTs, suggested that antioxidant vitamin supplementation was not associated with a reduced risk of cataracts which is consistent with our findings (158).

NPHS is not a nutrition-focused survey. The information regarding FV consumption was not detailed documented in NPHS. For example, the food items only limited to the following: fruit juices (orange, grapefruit, or tomato etc.), fruit, green salad, potatoes, carrots, and other vegetables. In addition, FV intake was only measured as frequency, and the portion size was not recorded, which may lead to inaccurate measurement of FV intake.

Residual confounding may still exist in our studies, even though we have included comprehensive confounders. For example, UV exposure and health care encounter are potential confounders, but we could not control for in our study due to data availability. Ignorance of them and/or other unknown confounders could produce distorted estimation.

6.2 Strengths

Our study has several strengths. The use of a national population-based longitudinal study allowed us to follow participants up to 17 years. The longitudinal nature of this study enabled us to avoid reverse causality when engaging in a causal inference. Furthermore, recall bias was avoided in the longitudinal study, because the outcome of interest occurs after the exposures are assessed. The non-response rate was relatively low in NPHS, with a response rate ranging from 69.7% to 92.8%, which reduced the attrition bias and ensured a larger sample size to increase statistical power. The NPHS collected health information and related socio-demographic information, which allowed us to control for relevant potential confounders.

Censoring due to loss to follow-up or end of study is not uncommon for observational studies, which could introduce selection bias and pose a threat to the internal validity of estimates (159). We used Cox proportional hazards model to incorporate censoring into the analysis. Previous studies have a varied definition of alcohol use and FV intake. We treated both alcohol and FV intake as continuous variables, which reduced the inaccurate estimation and loss of power that categorization could lead to (131).

Additionally, the Cox proportional hazards model allowed us to include time-varying covariates. For example, we computed the cumulative value of alcohol use and FV consumption and updated it up to the start of each follow-up, which enables us to estimate the long-term and cumulative effect of exposure. Moreover, we examined the effect modification by sex and age on the association of alcohol and FV consumption with cataracts.

Finally, we conducted multiple imputations using fully conditional specification (FCS) to address miss data in our study. The pooled parameters estimated from 30 imputed data sets provided evidence of robustness for our primary analysis.

6.3 Limitations

There are several limitations to our study. Due to the data availability of NPHS, information on FV consumption was only collected from Cycle 5, which limited our maximum follow-up time to 8 years for our second objective. We were unable to control for FV consumption in our first study objective. However, the sensitivity analysis using subset of the first cohort showed our estimation is still robust after accounting for different length of follow-up time and including FV in the model.

Another limitation was that cataract ascertainment was based on a self-reported professional-diagnosed cataract, and therefore, misclassification of cataract status could happen. As we used data from NPHS, which is a longitudinal population-based health survey rather than a cataract-focused study, misclassification is expected to be non-differential by our exposure of interest (e.g. alcohol use and FV intake). As a result, misclassification will bias the estimation towards the null. Furthermore, we could not examine the effect of exposures on the risks of specific types of cataracts. We also could not assess the effect of different types of alcoholic beverages on

cataracts. Moreover, as NPHS is not a nutritional study, only a limited number of food items were collected in the questionnaire. In addition, FV was only measured in frequency, and the portion size was not collected. Measurement error may have occurred due to insufficient information on FV intake.

Finally, our study was subject to weaknesses inherent in any observational studies. Residual confounding cannot be eliminated and remains a possible explanation of our estimates.

6.4 Conclusion

The present study examined the effects of alcohol use and FV intake on the risk of cataracts. Results showed neither alcohol use nor FV intake was statistically significantly associated with cataracts. No evidence was found for the heterogeneous effects by sex or age of the association.

More high-quality observational studies with linkage of administrative data are needed to ensure more detailed and accurate data collection to comprehensively measure potential confounders as well as outcomes, as controlled clinical trials are not feasible to investigate the effect of alcohol use. In addition, as alcohol use and FV intake was only collected during every visit. It is important to apply validated food diaries so that more accurate and detailed use information are collected.

In terms of the association of FV intake with cataracts, controlled intervention studies are warranted to address the effect of the consumption of different types of foods on the risk of cataracts.

In conclusion, our findings suggest that neither alcohol use nor FV intake is associated with cataracts, which contributes to the existing body of literature.

References

1. Wright KW (Kenneth W, editor. Textbook of ophthalmology / [edited by] Kenneth W. Wright ; section editors, Edward G. Buckley ... [et al.] ; consultants, Kenneth C. Chern ... [et al.] ; illustrator, Timothy C. Hengst. Baltimore: Williams & Wilkins; 1997.

2. Bowling B (Bradley). Kanski's clinical ophthalmology : a systematic approach / Brad Bowling. Eighth edi. Kanski JJ, editor. Clinical ophthalmology. Edinburgh: Elsevier; 2016.

3. Chan WH, Biswas S, Ashworth JL, Lloyd IC. Congenital and infantile cataract: aetiology and management. Eur J Pediatr [Internet]. 2012;171(4):625–30. Available from: http://europepmc.org/abstract/MED/22383071

4. Floud S, Kuper H, Reeves GK, Beral V, Green J, S. F, et al. Risk Factors for Cataracts Treated Surgically in Postmenopausal Women. Ophthalmology [Internet]. 2016 Aug;123(8):1704–10. Available from: http://www.elsevier.com/locate/ophtha

5. Storey, P., Munoz, B., Friedman, D., & West S. Racial differences in lens opacity incidence and progression: The salisbury eye evaluation (SEE) study. Investig Ophthalmol Vis Sci [Internet]. 2013 Apr;54(4):3010–8. Available from: http://www.iovs.org/content/54/4/3010.full.pdf+html

6. Bastawrous A, Mathenge W, Nkurikiye J, Wing K, Rono H, Gichangi M, et al. Incidence of Visually Impairing Cataracts among Older Adults in Kenya. JAMA Netw Open [Internet]. 2019 Jun;2(6):e196354. Available from: https://jamanetwork.com/journals/jamanetworkopen

7. Wu, R., Wang, J. J., Mitchell, P., Lamoureux, E. L., Zheng, Y., Rochtchina, E., ... & Wong TY. Smoking, socioeconomic factors, and age-related cataract: The Singapore Malay Eye study. Arch Ophthalmol (Chicago, Ill 1960). 2010 Aug;128(8):1029–35.

8. Varma R, Torres M. Prevalence of lens opacities in Latinos: the Los Angeles Latino Eye Study. Ophthalmology. 2004 Aug;111(8):1449–56.

9. National Eye Institute. Cataract Data and Statistics [Internet]. 2017. Available from: https://www.nei.nih.gov/learn-about-eye-health/resources-for-health-educators/eye-health-

data-and-statistics/cataract-data-and-statistics

10. Statistics Canada. Healthy aging indicators [Internet]. 2013. Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310046601

11. Hashemi H, Pakzad R, Yekta A, Aghamirsalim M, Pakbin M, Ramin S, et al. Global and regional prevalence of age-related cataract: a comprehensive systematic review and meta-analysis. Eye (Lond). 2020 Aug;34(8):1357–70.

12. Khairallah M, Kahloun R, Bourne R, Limburg H, Flaxman SR, Jonas JB, et al. Number of People Blind or Visually Impaired by Cataract Worldwide and in World Regions, 1990 to 2010. Invest Ophthalmol Vis Sci. 2015 Oct;56(11):6762–9.

13. Bourne RRA, Flaxman SR, Braithwaite T, Cicinelli M V, Das A, Jonas JB, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob Heal [Internet]. 2017;5(9):e888–97. Available from: http://www.sciencedirect.com/science/article/pii/S2214109X17302930

14. Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. Br J Ophthalmol [Internet]. 2012 May 1;96(5):614 LP – 618. Available from: http://bjo.bmj.com/content/96/5/614.abstract

15. Chew M, Chiang PP-C, Zheng Y, Lavanya R, Wu R, Saw SM, et al. The Impact of Cataract, Cataract Types, and Cataract Grades on Vision-Specific Functioning Using Rasch Analysis. Am J Ophthalmol [Internet]. 2012;154(1):29-38.e2. Available from: http://www.sciencedirect.com/science/article/pii/S0002939412000840

16. Polack S, Eusebio C, Fletcher A, Foster A, Kuper H. Visual Impairment from Cataract and Health Related Quality of Life: Results from a Case-Control Study in the Philippines. Ophthalmic Epidemiol [Internet]. 2010 Jun 1;17(3):152–9. Available from: https://doi.org/10.3109/09286581003731536

17. Klein, B. E., Moss, S. E., Klein, R., Lee, K. E., & Cruickshanks KJ. Associations of visual function with physical outcomes and limitations 5 years later in an older population: The Beaver Dam Eye Study. Ophthalmology [Internet]. 2003;110(4):644–50. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-

0037397018&doi=10.1016%2FS0161-6420%2802%2901935-8&partnerID=40&md5=56a223a808fd39c35ccda43869cdb276

18. Thomas BJ, Sanders DS, Oliva MS, Orrs MS, Glick P, Ruit S, et al. Blindness, cataract surgery and mortality in Ethiopia. Br J Ophthalmol. 2016 Sep;100(9):1157–62.

19. Bhattacharya R, Shen C, Sambamoorthi U. Excess risk of chronic physical conditions associated with depression and anxiety. BMC Psychiatry [Internet]. 2014;14:10. Available from: https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/docview/1491650884?accountid=15115

20. Almeida OP, Draper B, Pirkis J, Snowdon J, Lautenschlager NT, Byrne G, et al. Anxiety, depression, and comorbid anxiety and depression: Risk factors and outcome over two years. Int Psychogeriatrics [Internet]. 2012 Oct;24(10):1622–32. Available from: https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/docview/1081623338?accountid=15115

21. Choi HG, Lee MJ, Lee S-M. Visual impairment and risk of depression: A longitudinal follow-up study using a national sample cohort. Sci Rep [Internet]. 2018 Feb 1;8(1):2083. Available from: https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/docview/1993992204?accountid=15115

22. Grant A, Aubin M-J, Buhrmann R, Kergoat M-J, Freeman EE. Visual Impairment, Eye Disease, and the 3-year Incidence of Depressive Symptoms: The Canadian Longitudinal Study on Aging. Ophthalmic Epidemiol [Internet]. 2020 Sep 24;1–9. Available from: https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/docview/2446663974?accountid=15115

23. Zheng Y, Wu X, Lin X, Lin H. The Prevalence of Depression and Depressive Symptoms among Eye Disease Patients: A Systematic Review and Meta-analysis. Sci Rep [Internet]. 2017 Apr 12;7:46453. Available from: https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/docview/1887050943?accountid=15115

24. Rein DB, Zhang P, Wirth KE, Lee PP, Hoerger TJ, McCall N, et al. The Economic Burden of Major Adult Visual Disorders in the United States. Arch Ophthalmol [Internet]. 2006 Dec 1;124(12):1754–60. Available from:

https://doi.org/10.1001/archopht.124.12.1754

25. Taylor HR, Pezzullo ML, Keeffe JE. The economic impact and cost of visual impairment in Australia. Br J Ophthalmol [Internet]. 2006 Mar 1;90(3):272 LP – 275. Available from: http://bjo.bmj.com/content/90/3/272.abstract

26. Cruess AF, Gordon KD, Bellan L, Mitchell S, Pezzullo ML. The cost of vision loss in Canada. 2. Results. Can J Ophthalmol. 2011 Aug;46(4):315–8.

27. Congdon N, Vingerling JR, Klein BEK, West S, Friedman DS, Kempen J, et al. Prevalence of cataract and pseudophakia/aphakia among adults in the United States. Arch Ophthalmol (Chicago, Ill 1960). 2004 Apr;122(4):487–94.

28. Klein R, Klein BEK. The prevalence of age-related eye diseases and visual impairment in aging: current estimates. Invest Ophthalmol Vis Sci. 2013 Dec;54(14):ORSF5–13.

29. World Population Prospects: the 2019 Revision [Internet]. new york; 2019. Available from: https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf

30. Klein BE, Klein R, Linton KL. Prevalence of age-related lens opacities in a population. The Beaver Dam Eye Study. Ophthalmology. 1992 Apr;99(4):546–52.

31. Carlsson B, Sjostrand J. Increased incidence of cataract extractions in women above 70 years of age. A population based study. Acta Ophthalmol Scand. 1996 Feb;74(1):64–8.

32. Mitchell P, Cumming RG, Attebo K, Panchapakesan J. Prevalence of cataract in Australia: the Blue Mountains eye study. Ophthalmology. 1997 Apr;104(4):581–8.

33. Lundstrom M, Stenevi U, Thorburn W. Gender and cataract surgery in Sweden 1992-1997. A retrospective observational study based on the Swedish National Cataract Register. Acta Ophthalmol Scand. 1999 Apr;77(2):204–8.

34. Ostberg A, Loth A, Gustafson D, Lindblom B. Skovde cataract study: I. Prevalence of lens opacities in a Swedish community. Ophthalmology. 2006 Jun;113(6):970–5.

35. Erol Tinaztepe O, Ay M, Eser E. Nuclear and Mitochondrial DNA of Age-Related Cataract Patients Are Susceptible to Oxidative Damage. Curr Eye Res. 2017 Apr;42(4):583–8.

36. Borrás C, Gambini J, Gómez-Cabrera MC, Sastre J, Pallardó F V, Mann GE, et al. 17β-oestradiol up-regulates longevity-related, antioxidant enzyme expression via the ERK1 and ERK2 [MAPK]/NFκB cascade. Aging Cell. 2005;4(3):113–8.

37. Razmara A, Duckles SP, Krause DN, Procaccio V. Estrogen suppresses brain mitochondrial oxidative stress in female and male rats. Brain Res. 2007;1176:71–81.

38. Wang Y, Hunt K, Nazareth I, Freemantle N, Petersen I. Do men consult less than women? An analysis of routinely collected UK general practice data. BMJ Open [Internet]. 2013 Aug 1;3(8):e003320. Available from: http://bmjopen.bmj.com/content/3/8/e003320.abstract

39. West SK, Munoz B, Schein OD, Duncan DD, Rubin GS. Racial differences in lens opacities: the Salisbury Eye Evaluation (SEE) project. Am J Epidemiol. 1998 Dec;148(11):1033–9.

40. Chua J, Koh JY, Tan AG, Zhao W, Lamoureux E, Mitchell P, et al. Ancestry, Socioeconomic Status, and Age-Related Cataract in Asians: The Singapore Epidemiology of Eye Diseases Study. Ophthalmology. 2015 Nov;122(11):2169–78.

41. Stäubli A, Capatina N, Fuhrer Y, Munier FL, Labs S, Schorderet DF, et al. Abnormal creatine transport of mutations in monocarboxylate transporter 12 (MCT12) found in patients with age-related cataract can be partially rescued by exogenous chaperone CD147. Hum Mol Genet [Internet]. 2017 Aug 13;26(21):4203–14. Available from: https://doi.org/10.1093/hmg/ddx310

42. Klein BE, Klein R, Lee KE MS, K.E. L. Socioeconomic and lifestyle factors and the 10-year incidence of age-related cataracts. Am J Ophthalmol [Internet]. 2003 Sep;136(3):506–12. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed8&NEWS=N&AN=37101248

43. Hruba D. [Smoking and breast cancer]. Klin Onkol. 2013;26(6):389–93.

44. Jassem E, Szymanowska A, Sieminska A, Jassem J. [Smoking and lung cancer]. Pneumonol Alergol Pol. 2009;77(5):469–73.

45. Laniado-Laborín R. Smoking and chronic obstructive pulmonary disease (COPD). Parallel epidemics of the 21 century. Int J Environ Res Public Health [Internet]. 2009/01/09. 2009 Jan;6(1):209–24. Available from: https://www.ncbi.nlm.nih.gov/pubmed/19440278

46. Buist AS, Vollmer WM, McBurnie MA. Worldwide burden of COPD in high- and low-income countries. Part I. The burden of obstructive lung disease (BOLD) initiative. Int J Tuberc Lung Dis. 2008 Jul;12(7):703–8.

47. Waters D, Lespe´ rance J, Gladstone P, Boccuzzi SJ, Cook T, Hudgin R, et al. Effects of cigarette smoking on the angiographic evolution of coronary atherosclerosis: a Canadian Coronary Atherosclerosis Intervention Trial (CCAIT) substudy. Circulation. 1996;94(4):614–21.

48. Christen, W. G., Manson, J. E., Seddon, J. M., Glynn, R. J., Buring, J. E., Rosner, B., & Hennekens CH. A prospective study of cigarette smoking and risk of cataract in men. JAMA-JOURNAL Am Med Assoc. 1992 Aug;268(8):989–93.

49. HANKINSON SE, WILLETT WC, COLDITZ GA, SEDDON JM, ROSNER B, SPEIZER FE, et al. A PROSPECTIVE-STUDY OF CIGARETTE-SMOKING AND RISK OF CATARACT-SURGERY IN WOMEN. JAMA-JOURNAL Am Med Assoc. 1992 Aug;268(8):994–8.

50. Hiller R, Podgor MJ, Sperduto RD, Nowroozi L, Wilson PW, D'Agostino RB, et al. A longitudinal study of body mass index and lens opacities. The Framingham Studies. Ophthalmology. 1998 Jul;105(7):1244–50.

51. Kuang T-M, Tsai S-Y, Liu CJ-L, Ko Y-C, Lee S-M, Chou P. Seven-year incidence of age-related cataracts among an elderly Chinese population in Shihpai, Taiwan: The Shihpai Eye Study. Invest Ophthalmol Vis Sci. 2013 Sep;54(9):6409–15.

52. Tan AG, Kifley A, Flood VM, Holliday EG, Scott RJ, Cumming RG, et al. Evaluating the associations between obesity and age-related cataract: a Mendelian randomization study. Am J Clin Nutr. 2019 Oct;110(4):969–76.

53. Williams PT. Walking and running are associated with similar reductions in cataract risk. Med Sci Sports Exerc [Internet]. 2013;45(6):1089–96. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed14&NEWS=N&A

N=52372874

54. Zheng Selin J, Orsini N, Ejdervik Lindblad B, Wolk A. Long-term physical activity and risk of age-related cataract: a population-based prospective study of male and female cohorts. Ophthalmology. 2015 Feb;122(2):274–80.

55. Jiang H, Wang L-N, Liu Y, Li M, Wu M, Yin Y, et al. Physical activity and risk of age-related cataract. Int J Ophthalmol. 2020;13(4):643–9.

56. Cullen AP. Photokeratitis and other phototoxic effects on the cornea and conjunctiva. Int J Toxicol. 2002;21(6):455–64.

57. Roberts JE. Ultraviolet Radiation as a Risk Factor for Cataract and Macular Degeneration. Eye Contact Lens [Internet]. 2011;37(4). Available from: https://journals.lww.com/claojournal/Fulltext/2011/07000/Ultraviolet_Radiation_as_a_Risk_Factor_for.11.aspx

58. Kleiman NJ, Wang RR, Spector A. Ultraviolet light induced DNA damage and repair in bovine lens epithelial cells. Curr Eye Res. 1990 Dec;9(12):1185–93.

59. Zhang J, Yan H, Lofgren S, Tian X, Lou MF. Ultraviolet radiation-induced cataract in mice: the effect of age and the potential biochemical mechanism. Invest Ophthalmol Vis Sci. 2012 Oct;53(11):7276–85.

60. Cruickshanks KJ, Klein BE, Klein R. Ultraviolet light exposure and lens opacities: the Beaver Dam Eye Study. Am J Public Health. 1992 Dec;82(12):1658–62.

61. Delcourt C, Carriere I, Ponton-Sanchez A, Lacroux A, Covacho MJ, Papoz L. Light exposure and the risk of cortical, nuclear, and posterior subcapsular cataracts: the Pathologies Oculaires Liees a l'Age (POLA) study. Arch Ophthalmol (Chicago, Ill 1960). 2000 Mar;118(3):385–92.

62. Rice JB, White AG, Scarpati LM, Wan G, Nelson WW. Long-term Systemic Corticosteroid Exposure: A Systematic Literature Review. Clin Ther. 2017 Nov;39(11):2216–29.

63. Kačmař J, Cholevík D. Corticosteroid Induced Posterior Subcapsular Cataract. Ces a Slov Oftalmol Cas Ces Oftalmol Spol a Slov Oftalmol Spol. 2019;74(6):226–32.

64. Fel A, Aslangul E, Le Jeunne C. [Eye and corticosteroid's use]. Presse Med. 2012 Apr;41(4):414–21.

65. Wang JJ, Rochtchina E, Tan AG, Cumming RG, Leeder SR, Mitchell P. Use of inhaled and oral corticosteroids and the long-term risk of cataract. Ophthalmology. 2009 Apr;116(4):652–7.

66. Smeeth L, Boulis M, Hubbard R, Fletcher AE. A population based case-control study of cataract and inhaled corticosteroids. Br J Ophthalmol. 2003 Oct;87(10):1247–51.

67. Skalka HW, Prchal JT. Effect of Corticosteroids on Cataract Formation. Arch Ophthalmol [Internet]. 1980 Oct 1;98(10):1773–7. Available from: https://doi.org/10.1001/archopht.1980.01020040625007

68. Sinha N, Kumar Dabla P. Oxidative stress and antioxidants in hypertension–a current review. Curr Hypertens Rev. 2015;11(2):132–42.

69. Shichi H. Cataract formation and prevention. Expert Opin Investig Drugs. 2004 Jun;13(6):691–701.

70. Halliwell B. Antioxidants in human health and disease. Annu Rev Nutr. 1996;16(1):33–50.

71. Almeida IMC, Barreira JCM, Oliveira MBPP, Ferreira ICFR. Dietary antioxidant supplements: benefits of their combined use. Food Chem Toxicol. 2011;49(12):3232–7.

72. Zheng Selin J, Rautiainen S, Lindblad BE, Morgenstern R, Wolk A, J. ZS, et al. High-dose supplements of vitamins C and E, low-dose multivitamins, and the risk of age-related cataract: A population-based prospective cohort study of men. Am J Epidemiol [Internet]. 2013 Mar;177(6):548–55. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed14&NEWS=N&AN=368512148

73. Jiang H, Yin Y, Wu C-R, Liu Y, Guo F, Li M, et al. Dietary vitamin and carotenoid intake and risk of age-related cataract. Am J Clin Nutr. 2019 Jan;109(1):43–54.

74. Mukesh BN, Le A, Dimitrov PN, Ahmed S, Taylor HR, McCarty CA. Development of Cataract and Associated Risk Factors: The Visual Impairment Project. JAMA Ophthalmol

[Internet]. 2006 Jan 1;124(1):79–85. Available from: https://doi.org/10.1001/archopht.124.1.79

75. Younan C, Mitchell P, Cumming RG, Rochtchina E WJ. Myopia and incident cataract and cataract surgery: the blue mountains eye study. Vol. 43, Investigative ophthalmology & visual science. [St. Louis] : CV Mosby,; 2002. p. 3625.

76. Pan C-W, Cheng C-Y, Saw S-M, Wang JJ, Wong TY. Myopia and age-related cataract: a systematic review and meta-analysis. Am J Ophthalmol. 2013 Nov;156(5):1021-1033.e1.

77. Li L, Wan X, Zhao G. Meta-analysis of the risk of cataract in type 2 diabetes. BMC Ophthalmol. 2014 Jul;14:94.

78. Machan CM, Hrynchak PK, Irving EL. Age-related cataract is associated with type 2 diabetes and statin use. Optom Vis Sci. 2012 Aug;89(8):1165–71.

79. Olafsdottir E, Andersson DKG, Stefansson E. The prevalence of cataract in a population with and without type 2 diabetes mellitus. Acta Ophthalmol. 2012 Jun;90(4):334–40.

80. Foster PJ, Wong TY, Machin D, Johnson GJ, Seah SKL. Risk factors for nuclear, cortical and posterior subcapsular cataracts in the Chinese population of Singapore: the Tanjong Pagar Survey. Br J Ophthalmol. 2003 Sep;87(9):1112–20.

81. Jacques PF, Moeller SM, Hankinson SE, Chylack LTJ, Rogers G, Tung W, et al. Weight status, abdominal adiposity, diabetes, and early age-related lens opacities. Am J Clin Nutr. 2003 Sep;78(3):400–5.

82. Storey, P., Munoz, B., Friedman, D., & West S. Diabetes, fasting blood glucose and age-related cataract: The Blue Mountains Eye Study. Ophthalmic Epidemiol [Internet]. 2000 Jun;7(2):103–14. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed7&NEWS=N&AN=30495375

83. Delcourt C, Cristol JP, Tessier F, Leger CL, Michel F, Papoz L. Risk factors for cortical, nuclear, and posterior subcapsular cataracts: the POLA study. Pathologies Oculaires Liees a l'Age. Am J Epidemiol. 2000 Mar;151(5):497–504.

84. Leske MC, Wu SY, Hennis A, Connell AM, Hyman L, Schachat A. Diabetes,

hypertension, and central obesity as cataract risk factors in a black population. The Barbados Eye Study. Ophthalmology. 1999 Jan;106(1):35–41.

85. Klein BE, Klein R, Jensen SC, Linton KL. Hypertension and lens opacities from the Beaver Dam Eye Study. Am J Ophthalmol. 1995 May;119(5):640–6.

86. Schaumberg DA, Glynn RJ, Christen WG, Ajani UA, Sturmer T, Hennekens CH. A prospective study of blood pressure and risk of cataract in men. Ann Epidemiol. 2001 Feb;11(2):104–10.

87. Mylona I, Dermenoudi M, Ziakas N, Tsinopoulos I. Hypertension is the Prominent Risk Factor in Cataract Patients. Medicina (Kaunas). 2019 Aug;55(8).

88. Goodrich ME, Cumming RG, Mitchell P, Koutts J, Burnett L. Plasma fibrinogen and other cardiovascular disease risk factors and cataract. Ophthalmic Epidemiol. 1999 Dec;6(4):279–90.

89. Shiels A, Hejtmancik JF. Mutations and mechanisms in congenital and age-related cataracts. Exp Eye Res [Internet]. 2016/06/19. 2017 Mar;156:95–102. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27334249

90. Yang J, Luo J, Zhou P, Fan Q, Luo Y, Lu Y. Association of the ephreceptor tyrosinekinase-type A2 (EPHA2) gene polymorphism rs3754334 with age-related cataract risk: a meta-analysis. PLoS One. 2013;8(8):e71003.

91. Andersson ME, Zetterberg M, Tasa G, Seibt-Palmer M, Juronen E, Teesalu P, et al. Variability in the kinesin light chain 1 gene may influence risk of age-related cataract. Mol Vis. 2007 Jun;13:993–6.

92. Wu M, Zheng C, Yuan R-D, Sun M, Xu Y, Ye J. The link between apolipoprotein E, presenilin 1, and kinesin light chain 1 gene polymorphisms and age-related cortical cataracts in the Chinese population. Mol Vis. 2015;21:412–6.

93. Unal M, Guven M, Batar B, Ozaydin A, Sarici A, Devranoglu K. Polymorphisms of DNA repair genes XPD and XRCC1 and risk of cataract development. Exp Eye Res. 2007 Sep;85(3):328–34.

94. Nizami, AA. Gulani A. Cataract [Internet]. StatPearls Publishing. 2020. Available from:

https://www.ncbi.nlm.nih.gov/books/NBK539699/

95. Lasa MS, Podgor MJ, Datiles 3rd MB, Caruso RC, Magno B V. Glare sensitivity in early cataracts. Br J Ophthalmol [Internet]. 1993 Aug;77(8):489–91. Available from: https://www.ncbi.nlm.nih.gov/pubmed/8025045

96. Cologne GI for Q and E in HC. Cataracts: Overview [Internet]. InformedHealth.org. 2019. Available from: https://www.ncbi.nlm.nih.gov/books/NBK390302/

97. James B. Ophthalmology : lecture notes / Bruce James, Anthony Bron, Manoj V. Parulekar. Twelfth ed. Bron AJ, Parulekar M V, editors. Chichester, England: Wiley Blackwell; 2017.

98. Gali HE, Sella R, Afshari NA. Cataract grading systems: a review of past and present. Curr Opin Ophthalmol. 2019 Jan;30(1):13–8.

99. Allen D, Vasavada A. Cataract and surgery for cataract. BMJ [Internet]. 2006 Jul 15;333(7559):128–32. Available from: https://www.ncbi.nlm.nih.gov/pubmed/16840470

100. Lamoureux EL, Fenwick E, Pesudovs K, Tan D. The impact of cataract surgery on quality of life. Curr Opin Ophthalmol. 2011 Jan;22(1):19–27.

101. Tseng VL, Yu F, Lum F, Coleman AL. Cataract Surgery and Mortality in the United States Medicare Population. Ophthalmology. 2016 May;123(5):1019–26.

102. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009 Oct;62(10):1006–12.

103. Lindblad BE, Hakansson N, Philipson B, Wolk A. Alcohol consumption and risk of cataract extraction: a prospective cohort study of women. Ophthalmology. 2007 Apr;114(4):680–5.

104. Kuang, T. M., Tsai, S. Y., Liu, C. J. L., Ko, Y. C., Lee, S. M., & Chou P. Seven-year incidence of age-related cataracts among an elderly Chinese population in Shihpai, Taiwan: The Shihpai Eye Study. Investig Ophthalmol Vis Sci [Internet]. 2013;54(9):6409–15. Available from: http://www.iovs.org/content/54/9/6409.full.pdf

105. Kanthan GL, Mitchell P, Burlutsky G, Wang JJ, G.L. K, P. M, et al. Alcohol consumption and the long-term incidence of cataract and cataract surgery: The blue mountains eye study. Am J Ophthalmol [Internet]. 2010 Sep;150(3):434. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed11&NEWS=N&AN=50982390

106. Manson, J. E., Christen, W. G., Seddon, J. M., Glynn, R. J., & Hennekens CH. A prospective study of alcohol consumption and risk of cataract. Am J Prev Med [Internet]. 1994;10(3):156–61. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed5&NEWS=N&AN=24203516

107. Chasan-Taber, L., Willett, W. C., Seddon, J. M., Stampfer, M. J., Rosner, B., Colditz, G. A., ... & Hankinson SE, F.E. S, Chasan-Taber L, Willett WC, Seddon JM, Stampfer MJ, et al. A prospective study of alcohol consumption and cataract extraction among U.S. women. Ann Epidemiol [Internet]. 2000 Aug;10(6):347–53. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed7&NEWS=N&AN=30645788

108. Nam, G. E., Han, K., Ha, S. G., Han, B. D., Kim, D. H., Kim, Y. H., ... & Ko BJ. Relationship between socioeconomic and lifestyle factors and cataracts in Koreans: The Korea National Health and Nutrition Examination Survey 2008-2011. Eye [Internet]. 2015;29(7):913–20. Available from: http://www.nature.com/eye/index.html

109. Li Z, Xu K, Wu S, Sun Y, Song Z, Jin D, et al. Alcohol consumption and visual impairment in a rural Northern Chinese population. Ophthalmic Epidemiol. 2014 Dec;21(6):384–90.

110. Morris MS, Jacques PF, Hankinson SE, Chylack LTJ, Willett WC, Taylor A. Moderate alcoholic beverage intake and early nuclear and cortical lens opacities. Ophthalmic Epidemiol. 2004 Feb;11(1):53–65.

111. Tsai S-Y, Hsu W-M, Cheng C-Y, Liu J-H, Chou P, S.-Y. T, et al. Epidemiologic study of age-related cataracts among an elderly Chinese population in Shih-Pai, Taiwan. Ophthalmology [Internet]. 2003 Jun;110(6):1089–95. Available from:

http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed8&NEWS=N&AN=36676638

112. McCarty CA, Mukesh BN, Fu CL, Taylor HR. The epidemiology of cataract in Australia. Am J Ophthalmol. 1999 Oct;128(4):446–65.

113. Cumming RG, Mitchell P. Alcohol, smoking, and cataracts: the Blue Mountains Eye Study. Arch Ophthalmol (Chicago, Ill 1960). 1997 Oct;115(10):1296–303.

114. Singh, S., Pardhan, S., Kulothungan, V., Swaminathan, G., Ravichandran, J. S., Ganesan, S., ... & Raman R. The prevalence and risk factors for cataract in rural and urban India. Indian J Ophthalmol [Internet]. 2019 Apr;67(4):477–83. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emexa&NEWS=N&AN=626992428

115. Tavani A, Negri E LVC. Food and nutrient intake and risk of cataract. Ann Epidemiol [Internet]. 1996 Jan;6(1):41–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed6&NEWS=N&AN=26056112

116. Phillips, C. I., Clayton, R. M., Cuthbert, J., Qian, W., Donnelly, C. A., & Prescott RJ. Human cataract risk factors: Significance of abstention from, and high consumption of, ethanol (U-Curve) and non-significance of smoking. Ophthalmic Res [Internet]. 1996;28(4):237–47. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed6&NEWS=N&AN=26294397

117. Bastawrous A, Mathenge W, Nkurikiye J, Wing K, Rono H, Gichangi M, et al. Incidence of Visually Impairing Cataracts Among Older Adults in Kenya. JAMA Netw open. 2019 Jun;2(6):e196354.

118. Frøslie KF, Røislien J, Laake P, Henriksen T, Qvigstad E, Veierød MB. Categorisation of continuous exposure variables revisited. A response to the Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study. BMC Med Res Methodol [Internet]. 2010;10(1):103. Available from: https://doi.org/10.1186/1471-2288-10-103

119. Vinson JA. Oxidative stress in cataracts. Pathophysiol Off J Int Soc Pathophysiol. 2006

Aug;13(3):151–62.

120. Adachi, S., Sawada, N., Yuki, K., Uchino, M., Iwasaki, M., Tsubota, K., & Tsugane S. Intake of vegetables and fruits and the risk of cataract incidence in a Japanese population: the Japan Public Health Center-based Prospective Study. J Epidemiol [Internet]. 2019; Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emexb&NEWS=N&AN=630218452

121. Rautiainen S, Lindblad BE, Morgenstern R, Wolk A, S. R, B.E. L, et al. Total antioxidant capacity of the diet and risk of age-related cataract: A population-based prospective cohort of women. JAMA Ophthalmol [Internet]. 2014 Mar;132(3):247–52. Available from: http://archopht.jamanetwork.com/article.aspx?articleid=1793735

122. Christen WG, Liu S, Schaumberg DA, Buring JE, W.G. C, S. L, et al. Fruit and vegetable intake and the risk of cataract in women. Am J Clin Nutr [Internet]. 2005 Jun;81(6):1417–22. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed9&NEWS=N&AN=41710247

123. Chasan-Taber, L., Willett, W. C., Seddon, J. M., Stampfer, M. J., Rosner, B., Colditz, G. A., ... & Hankinson SE. A prospective study of carotenoid and vitamin A intakes and risk of cataract extraction in US women. Am J Clin Nutr [Internet]. 1999 Oct;70(4):509–16. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed6&NEWS=N&AN=29463445

124. Theodoropoulou S, Samoli E, Theodossiadis PG, Papathanassiou M, Lagiou A, Lagiou P, et al. Diet and cataract: a case-control study. Int Ophthalmol. 2014 Feb;34(1):59–68.

125. Ghanavati M, Behrooz M, Rashidkhani B, Ashtray-Larky D, Zameni SD, Alipour M, et al. Healthy eating index in patients with cataract: A case-control study. Iran Red Crescent Med J [Internet]. 2015 Oct;17(10):e22490. Available from: http://ircmj.com/47601.pdf

126. Lu, Z. Q., & Yan J. Influence of fruit and vegetable intake on age-related cataract. Int J Ophthalmol [Internet]. 2012;12(1):58–61. Available from:

http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed13&NEWS=N&AN=364784649

127. Ojofeitimi, E. O., Adelekan, D. A., Adeoye, A., Ogungbe, T. G., Imoru, A. O., & Oduah EC. Dietary and lifestyle patterns in the aetiology of cataracts in Nigerian patients. Nutr Heal (Berkhamsted, Hertfordshire) [Internet]. 1999;13(2):61–8. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed6&NEWS=N&AN=129460635

128. Pastor-Valero M. Fruit and vegetable intake and vitamins C and E are associated with a reduced prevalence of cataract in a Spanish Mediterranean population. BMC Ophthalmol. 2013 Oct;13:52.

129. Moise MM, Benjamin L-M, Doris TM, Dalida KN, Augustin NO. Role of Mediterranean diet, tropical vegetables rich in antioxidants, and sunlight exposure in blindness, cataract and glaucoma among African type 2 diabetics. Int J Ophthalmol. 2012 Apr;5(2):231–7.

130. Cumming, R. G., Mitchell, P., & Smith W. Diet and cataract: the blue mountains eye study. Ophthalmology [Internet]. 2000 Mar;107(3):450–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed7&NEWS=N&AN=30191364

131. Bennette C, Vickers A. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. BMC Med Res Methodol [Internet]. 2012;12(1):21. Available from: https://doi.org/10.1186/1471-2288-12-21

132. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet (London, England). 2007 Oct;370(9596):1453–7.

133. Satistics Canada. National Population Health Survey [Internet]. 2012. Available from: https://www150.statcan.gc.ca/n1/en/pub/82f0068x/4198604-eng.pdf?st=ILtk9bC7

134. Statistics Canada. Documentation for the Derived Variables and the Constant Longitudinal Variables: Cycle1-Cycle9 [Internet]. Statistics Canada. 2012. Available from: https://www23.statcan.gc.ca/imdb-bmdi/pub/document/3225_D10_T9_V3-eng.pdf

135. Bland JM, Altman DG. Survival probabilities (the Kaplan-Meier method). Bmj. 1998;317(7172):1572–80.

136. van Walraven C, Davis D, Forster AJ, Wells GA. Time-dependent bias was common in survival analyses published in leading clinical journals. J Clin Epidemiol. 2004 Jul;57(7):672–82.

137. Kleinbaum DG. Survival analysis : a self-learning text / David G. Kleinbaum, Mitchel Klein. 2nd ed. Klein M, editor. New York, NY: Springer; 2005. (Statistics for biology and health).

138. Allison PD. Survival Analysis Using SAS: A Practical Guide. Second Edi. Sas Institute; 2010.

139. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation. 2016;133(6):601–9.

140. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94(446):496–509.

141. Berglund P, Heeringa SG. Multiple imputation of missing data using SAS. SAS Institute; 2014.

142. Little RJA, Rubin DB. Statistical analysis with missing data. Vol. 793. John Wiley & Sons; 2019.

143. Tsiampalis T, Panagiotakos D. Missing-data analysis: Socio- demographic, clinical and lifestyle determinants of low response rate on self- reported psychological and nutrition related multi- item instruments in the context of the ATTICA epidemiological study. BMC Med Res Methodol. 2020 Jun 8;20.

144. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med. 1999 Mar;18(6):681–94.

145. Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. Surv Methodol. 2001;27(1):85–96.

146. White IR, Royston P. Imputing missing covariate values for the Cox model. Stat Med [Internet]. 2009 Jul 10;28(15):1982–98. Available from: https://pubmed.ncbi.nlm.nih.gov/19452569

147. Efron B. The efficiency of Cox's likelihood function for censored data. J Am Stat Assoc. 1977;72(359):557–65.

148. Hertz-Picciotto I, Rockhill B. Validity and efficiency of approximation methods for tied survival times in Cox regression. Biometrics. 1997 Sep;53(3):1151–6.

149. Vittinghoff E, Glidden D V, McCulloch CE, Shiboski SC. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. 2. Aufl. New York, NY: Springer-Verlag; 2012. (Statistics for biology and health).

150. Lindblad, B. E., Håkansson, N., Philipson, B., & Wolk A. Alcohol Consumption and Risk of Cataract Extraction. A Prospective Cohort Study of Women. Ophthalmology [Internet]. 2007;114(4):680–5. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed10&NEWS=N&AN=46483681

151. Wang, W., & Zhang X. Alcohol intake and the risk of age-related cataracts: A meta-analysis of prospective cohort studies. PLoS One [Internet]. 2014;9(9):e107638. Available from: http://www.plosone.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0107820&representation=PDF

152. Gong Y, Feng K, Yan N, Xu Y, Pan C-W. Different amounts of alcohol consumption and cataract: a meta-analysis. Optom Vis Sci. 2015 Apr;92(4):471–9.

153. MacLennan PA, McGwin GJ, Searcey K, Owsley C. Medical record validation of self-reported eye diseases and eye care utilization among older adults. Curr Eye Res. 2013 Jan;38(1):1–8.

154. Delgado-Rodríguez M, Llorca J. Bias. J Epidemiol Community Health [Internet]. 2004 Aug 1;58(8):635 LP – 641. Available from: http://jech.bmj.com/content/58/8/635.abstract

155. Fewell Z, Davey Smith G, Sterne JAC. The impact of residual and unmeasured

confounding in epidemiologic studies: a simulation study. Am J Epidemiol. 2007 Sep;166(6):646–55.

156. Aziz M, Karboune S. Natural antimicrobial/antioxidant agents in meat and poultry products as well as fruits and vegetables: A review. Crit Rev Food Sci Nutr. 2018 Feb;58(3):486–511.

157. Serafini M, Peluso I. Functional Foods for Health: The Interrelated Antioxidant and Anti-Inflammatory Role of Fruits, Vegetables, Herbs, Spices and Cocoa in Humans. Curr Pharm Des. 2016;22(44):6701–15.

158. Mathew MC, Ervin A-M, Tao J, Davis RM. Antioxidant vitamin supplementation for preventing and slowing the progression of age-related cataract. Cochrane database Syst Rev. 2012 Jun;6(6):CD004567.

159. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004;615–25.

Appendices

Appendix A. Summary of previous studies on alcohol intake and cataracts

Author and Publication Year	Study Name	Study Design	Location	Population	Exposure Measurement	Outcome Measurement	Follow-up Period	Result
Singh et al. (2019)	NA	population-based cross-sectional studies	India	people aged 60 years and above(n=4,331)	alcohol drinking vs never drinking	Lens Opacification Classification System III	NA	No association between alcohol use and any cataract
Bastawrous et al. (2019)	NA	cohort studies	Kenya	people of East African ethnicity aged 50 years or older (n=4364)	current drinker and former drinker vs never drinker	visually impairing cataracts assessed by a simplified cataract grading system	Six years	With an RR of 1.4 (95%CI, 1.1-1.8) in current alcohol drinkers, compared with never drinkers, former drinkers were not at an increased risk (RR, 1.1; 95%CI, 0.9-1.3)
Floud et al. (2016)	The Million Women Study	Population-based cohort study	United Kingdom	postmenopausal women (n=1,312,051)	alcohol intake was measured by units and was categorized as <2, 2–14, ≥15 units per week (one unit was defined as 10g alcohol).	cataract surgery confirmed by medical records	an average of 11 years of follow-up	Alcohol consumption of 2 units or more per week was weakly associated with a slightly lower risk of cataracts surgery compared with consumption of less than 2 units per week, with RR of 2-14 units being 0.93 (95%CI, 0.91-0.94) and RR of ≥15units being 0.92 (95%CI, 0.89-0.96)
Nam et al. (2015)	The Korea National Health and Nutrition Examination Survey	Cross-sectional study	South Korea	people aged 40 years and above(n=15,866)	Participants were categorized as average alcohol intake: non-drinker, light to moderate drinker (1–30 g per day), and heavy drinker (>30 g per day)	Lens Opacification Classification System III	NA	No association between alcohol consumption and any cataract

Li et al. (2014)	NA	Cross-sectional study	China	people aged from 18-94 years (n=8445)	The frequency of alcohol intake was categorized as 0, 1-3,4-6, 7 days/week. Quantity of alcohol intake was classified as abstainers, former drinkers and current drinkers. Current drinkers were categorized as ≤1 drinks, 1-14 drinks and ≥14 drinks/week.	Diagnosed by ophthalmologist	NA	Increasing alcohol consumption was associated with decreased prevalence of cataract among those with light and moderate alcohol consumption(OR 0.9, 95%CI 0.4-2.4), whereas higher alcohol intakes were associated with an increased prevalence of cataract(OR 1.2, 95%CI 0.4-3.6)
Storey et al. (2013)	The Salisbury Eye Evaluation (SEE) Study	Population-based cohort study	United States	people aged from 65-84 years (n=2520)	current drinker and former drinker vs never drinker	Wilmer grading scheme	Two years	No association of nuclear opacity incidence with alcohol use with an OR of past and current alcohol use being 0.83(95% CI 0.50-1.37) and 0.90(95% CI 0.91-1.77). No association of cortical opacity incidence with alcohol use with an OR of past and current alcohol use being 1.45(95% CI 0.74-2.86) and 1.33 (95% CI 0.70-2.49).
Kuang et al. (2013)	The Shihpai Eye Study	cohort study	China	people aged 65 years and above (n=1361)	Alcohol consumption was limited to wine and hard alcohol. Participants were classified as drinker and non-drinker	Lens Opacification Classification System III	Seven years	Cortical cataract was not associated with alcohol use with an RR of 0.53 (95%CI, 0.26-1.07). No results were presented for nuclear and posterior subcapsular cataract.
Wu et al. (2010)	The Singapore Malay Eye Study	population-based cross-sectional studies	Singapore	people aged from 40-80 years (n=3280)	alcohol drinking vs never drinking	Wisconsin Cataract Grading System	NA	No association between alcohol consumption and any particular type of cataract or any cataract.
Kanthan et al. (2010)	The Blue Mountains Eye Study	Population-based prospectiv	Australia	people aged 49 years and above (n=3654)	Drinking patterns were classified into four categories	Wisconsin Cataract Grading System	5 years and 10 years	No associations were found between alcohol consumption and the incidence of any cataract subtypes. Abstainer and heavy drinkers (>2 drinks per day) were more likely to have cataract

		e cohort study			(none, ≤1, >1 but ≤2, and >2 drinks per day).			surgery with OR of 2.13(CI, 1.18-3.84) and 2.10 (CI 1.21-3.65) respectively, compared to those who drank 1 to 2 standard drinks of total alcohol per day
Lindblad et al. (2007)	Swedish Mammography Cohort	Population-based prospective cohort study	Sweden	Women aged from 49-83 years (n=34,713)	Drinkers were categorized as never drinkers, past drinkers and current drinkers. Current drinkers were then classified as (<6, 6–13, >13–20, >20–30 >30 g/day).	The Swedish National Cataract Register	84 months	The risk for cataract extraction increased with increasing alcohol use, with RR rising from 1.11 to 1.29.
Morris et al. (2004)	NA	Cross-sectional study	United States	Women aged from 53-74 years (n=556)	The frequency and number of drinks containing liquor, wine and beer were recorded. Grams of alcohol consumed per day was calculated and included the model.	Lens Opacification Classification System III	NA	The odds of a nuclear opacity increased by 30%(OR=1.3, 95%CI:1.10-1.54) per 10g increase in average daily intake of alcohol from all sources combined. Average daily alcohol use was not significantly associated with cortical opacity.
Tsai et al. (2003)	NA	population-based cross-sectional studies	China	people aged 65 years and above(n=1,361)	Alcohol use (yes vs no)	Lens Opacification Classification System III	NA	No association between alcohol consumption and any particular type of cataract. Age and gender had a significant interaction effect on posterior subcapsular cataracts, but not cortical and nuclear cataracts.
Klein et al. (2003)	The Beaver Dam Eye Study	Population-based prospective cohort study	United States	People aged from 43-86 years (n=4,926)	Each participant was asked for the consumption of beer, liquor and wine in the last year. The number of each alcoholic beverage they used each week was asked. The response was then converted to grams of alcohol.	Wisconsin Cataract Grading system	Ten years	A significant association between alcohol use and incidence of cataracts was not found. Alcohol use and smoking did not present a significant interaction effect on cataracts.

					Alcohol use was then classified as (none, 0-39, 40-99, 6–13, 100–334, 335+ g/week).			
Chasan-taber et al. (2000)	The Nurses' Health Study	Cohort study	United States	Nurses aged from 30-55 years(n=77,466)	Average daily alcohol use was derived from a semiquantitative food frequency questionnaire. In the analysis, alcohol intake was then classified as (almost never, ≤4.9, 5.0-14.9, 15.0–24.9, ≥25 g/day).	Self-reported and confirmed by medical records	12 years	No association was reported for total cataract and any cataract subtypes requiring extraction and increased alcohol use.
McCarty et al. (2000)	NA	Cross-sectional study	Australia	people aged 40 years and above(n=5,147)	alcohol use was classified as (none, ≤1, ≤2, ≤4, >4 drinks/day). Alcohol status was also categorized as current drinker, never drinker, past drinker.	Wilmer grading scheme	NA	alcohol use was associated with cortical and nuclear cataract but not posterior subcapsular cataract.
Cumming et al. (1997)	The Blue Mountains Eye Study	population-based, cross-sectional study	Australia	People aged from 49-97 years(n=3,654)	Alcohol intake was assessed by the frequency of drinking alcohol (days per week), the usual number of drinks on a day when alcohol was consumed. Alcohol use was classified as (none, <1, 1-3, and ≥4 drinks/day).	Wisconsin Cataract Grading System	NA	Heavy alcohol use (≥4 drinks/day) was only associated with nuclear cataract with an OR of 2.1 (CI, 1.1-4.3), whereas the association between it and cortical and subcapsular cataracts was not significant. Moderate alcohol use(1-3 drinks/day) was associated with a lower odd of cortical cataracts.
Tavani et al. (1996)	NA	Case control study	Italy	People aged from 21-80 years(n=913)	Alcohol intake was classified as (none, <3, 3-5, 5-8, and ≥8 drinks/day).	Diagnosed by ophthalmologist	NA	Total alcohol was not associated with cataract extraction.

Phillips et al. (1996)	NA	Case control study	Scotland	Stringently matched cataract-control pairs were included from the same hospital(n=1, 848)	Not stated	Medical record	NA	light and infrequent' and 'light and frequent' use of alcohol were associated with a lower risk of cataract than were life abstainer and 'occasional' consumption; the prevalence of cataract rose with further increases in consumption, suggesting a U-shaped curve.
Manson et al. (1994)	Physicians' Health Study	Prospective cohort study	United States	Male aged from 40-84 years (n=17,824)	Not stated	Self-report and confirmed by medical record	Five years	daily consumers of alcohol had a relative risk (RR) of cataract of 1.31 (95% CI= 0.95, 1.81), compared with participants who used alcohol less than once per month.

Abbreviations: CI=Confidence interval, RR=Relative risk, OR=Odds ratio, N/A=Not applicable

Appendix B. Summary of previous studies on FV intake and cataracts

Author and Publication Year	Study Name	Study Design	Location	Population	Exposure Measurement	Outcome Measurement	Follow-up Period	Result
Adachi et al. (2019)	Japan Public Health Center-based Prospective (JPHC) Study	Cohort study	Japan	People aged from 45-74 years (n=71,720)	Food frequency questionnaire (FFQ)	Self-reported	Five years	In men, total vegetable and cruciferous vegetable consumption were inversely associated with cataracts, with ORs of 0.77(95% CI,0.59-1.01) and 0.74 (95% CI, 0.57-0.96), respectively. However, the women showed an opposite trend, with an OR of 1.28 (95% CI, 1.06-1.53). There is no association between green and yellow vegetable and fruit intake with cataract.
Ghanavati et al. (2015)	NA	Case control study	Iran	97 cases and 198 controls	Healthy Eating Index(HEI) scores based on the Food frequency questionnaire (FFQ)	Diagnosed by ophthalmologist	NA	The HEI scores of vegetable were higher among healthy individuals than patients with cataracts (10 vs 7.8). The HEI scores of fruit were also higher among healthy individuals

								than patients with cataracts (9.8 vs 7.1).
Theodoropoulou et al.(2014)	NA	Case control study	Greece	314 cases and 314 frequency-matched controls aged from 45-85 years old	Semi-quantitative food-frequency questionnaire	Diagnosed by ophthalmologist	NA	The odds of having cataracts were decreased by 53% (OR=0.47, 95% CI:0.38, 0.59) and 47% (OR=0.53, 95% CI:0.39, 0.72) respectively for those consumed vegetables per 56.19 times/month and fruits per 109.41 times/month
Rautiainen et al. (2014)	Swedish Mammography Cohort(SMC)	Cohort study	Sweden	Women aged from 49 to 83 years old (n=30,607)	Total antioxidant capacity (TAC) based on the Food Frequency Questionnaire(FFQ)	linkage to registers in the study area	average seven years	The rate ratio was 0.87 (95%CI: 0.79-0.96; P for trend = .03) when comparing the highest quintile of the TAC of the diet with the lowest quintile
Pastor-Valero et al.(2013)	European Eye study (EUREYE)	Cross-sectional	Spain	People aged over 65 years (n=599)	Semi-quantitative food frequency questionnaire	Cataract or cataract extraction diagnosed by ophthalmologist	NA	Intake of combined fruit and vegetable was associated with the prevalence of cataract or cataract extraction after

								adjustment (P-trend=0.008)
Moise et al. (2012)	NA	Cross-sectional	Congo	Type 2 diabetic patients (n=500)	Harvard semi-quantitative Food Frequency Questionnaire(FFQ) adapted for Africa	NA	NA	regular intake of leafy vegetables was significantly associated with cataracts in patients with type 2 diabetes
Lu et al. (2012)	NA	Case control study	China	360 cases and 368 controls aged 45-85 years	semiquantitative food-frequency questionnaire (FFQ)	Cataract diagnosed by ophthalmologist	NA	The intake of vegetable and fruit reduced odd of having cataracts by 19% (OR=0.81, 95% CI: 0.67,0.97) and 19% (OR=0.81, 95% CI: 0.69,0.94) respectively
Christen et al. (2005)	Women's Health Study (WHS)	Prospective cohort study	United States	Female health professionals aged 45 years (n=35, 724)	validated semiquantitative food-frequency questionnaire (SFFQ)	Cataract or cataract extraction diagnosed by an ophthalmologist or confirmed by medical records	Ten years	the highest quintile of intake of fruit and vegetables(median intake=10.0 servings/day) contributed to reduced 170-15% reduced risks(RR=0.83, 95% CI: 0.70, 0.99)s of cataract compared with lowest quintile(median intake=2.6 servings/day)

Cumming et al. (2000)	NA	Cross-sectional	Australia	People aged 49 to 97 years(n=2900)	Semi-quantitative food frequency questionnaire (FFQ)	Cataract diagnosed by an ophthalmologist	NA	the associations between various vegetable and nuclear cataracts were not statistically significant
Ojofeitimi et al. (1999)	NA	Case control study	Nigeria	31 cases and 31 controls	Structured questionnaire	Cataract diagnosed by an ophthalmologist	NA	Higher numbers of controls than patients consumed fruits and vegetables
Chasan-Taber et al. (1999)	Nurses' Health Study	Cohort study	United States	Female nurses aged 45–71 years (n=50,461)	Food-frequency questionnaire	Cataract extraction via questionnaire and confirmed by an ophthalmologist	12 years	Intake of 2 times/week of spinach and other greens reduced the risk of cataract extraction by 18% compared with consumption of less than one time/month (RR=0.82, 95% CI: 0.68, 0.98).
Tavani et al. (1996)	NA	Case control study	Italy	207 cases and 706 controls	Food-frequency questionnaire	Cataract diagnosed by an ophthalmologist	NA	specific food items (Cruciferae, Spinach, Tomatoes and Melon) had an inverse association with the risk of cataracts

Abbreviations: CI=Confidence interval, RR=Relative risk, OR=Odds ratio, N/A=Not applicable

Appendix C. RDC research proposal

Title: The effect of Alcohol intake and Fruit and Vegetable Consumption on the Risk of Cataracts

Requesting access to the Western University RDC
Rationale and objectives of the study:

Background. A cataract is opacification of the lens, which is the leading cause of blindness globally (Bourne et al., 2017). Cataracts can be classified as age-related cataracts, pediatric cataracts, and cataracts secondary to other causes (Liu, Wilkins, Kim, Malyugin, & Mehta, 2017). A cataract is a multifactorial disease, and many risk factors have been identified, although evidence for some of the risk factors is inconsistent(Huang et al., 2015; B. E. K. Klein, Klein, Lee, & Meuer, 2003; R. Klein & Klein, 2013; Li, Wan, & Zhao, 2014; Pan, Cheng, Saw, Wang, & Wong, 2013; Poh, Mohamed Abdul, Lamoureux, Wong, & Sabanayagam, 2016; Richter, Choudhury, Torres, Azen, & Varma, 2012; Shahbazi, Studnicki, & Warner-Hillard, 2015; Storey, Munoz, Friedman, & West, 2013; Taylor et al., 1988; Tian et al., 2015; Wang & Zhang, 2014). Age is one of the common risk factors. The ageing population is growing at a remarkable rate. One in six people in the world is predicted to be over 65 by 2050, up from one in 11 in 2019(*World Population Prospects: the 2019 Revision*, 2019). Accordingly, the prevalence of cataracts increases with age, from 3% at 40-49 years old to 81% at 80 years and older(Varma & Torres, 2004). The population aged over 60 years suffering visual loss and blindness from age-related cataracts are predicted to be around 800 million in 2020, while this number was only 400 million in 2000(Foster, 2000). Therefore, the identification of modifiable factors is of great importance for public health care.

Rationale. Alcohol consumption is responsible for most common chronic conditions, such as cancer, cirrhosis of the liver, cardiovascular disease, stroke and diabetes(Bagnardi et al., 2015; Rehm et al., 2017). Several studies have been conducted regarding the association between alcohol consumption and cataracts. However, epidemiologic evidence on this relationship is inconsistent. Two cross-sectional studies reported that alcohol consumption was positively associated with cataracts(Cumming & Mitchell, 1997; Morris et al., 2004). However, the other cross-sectional studies did not find a significant association between alcohol intake and cataracts(McCarty, Mukesh, Fu, & Taylor, 1999; Singh et al., 2019; Tsai, Hsu, Cheng, Liu, & Chou, 2003). While four case-control studies found heavy alcohol consumption increases the risk of cataracts, one case-control study did not find evidence for the association between alcohol intake and cataract extraction(Echebiri, Odeigah, & Myers, 2010; Munoz, Tajchman, Bochow, & West, 1993; Phillips et al., 1996; Tavani, Negri, & La Vecchia, 1996; Ughade, Zodpey, & Khanolkar, 1998). A few prospective studies have also not reported a consistent association between alcohol drinking and cataract. Most prospective studies have not found a clear relationship between alcohol consumption and cataracts(Chasan-Taber et al., 2000; B. E. K.

Klein et al., 2003; Kuang et al., 2013; Lindblad, Hakansson, Philipson, & Wolk, 2007; Manson, Christen, Seddon, Glynn, & Hennekens, 1994; Storey et al., 2013), whereas the Blue Mountains Eye Study(Kanthan, Mitchell, Burlutsky, & Wang, 2010) reported a U-shaped association between alcohol intake and risk of cataracts. The most recent prospective study reported a borderline association between alcohol consumption of 2 units or more and a lower risk of cataracts treated surgically(Floud, Kuper, Reeves, Beral, & Green, 2016). The reasons for inconsistent finding are varying. Due to the nature of case-control studies, recall bias may distort the estimate of association. Temporality is a concern in terms of cross-sectional study design. More research with higher quality is needed to examine the association between alcohol consumption and cataracts. In addition, measurements of alcohol consumption are not consistent. Some studies only compared current drinker and former drinker with the ones who never drank. Although some studies categorize drink as light, moderate and heavy drink, the definition of each category is varying between studies. There is a need to conduct a study accounting for frequency, amount and drinking patterns of alcohol. Finally, some studies only adjusted limited confounders, which could also distort the association estimate. A study with adjusting comprehensive confounders is warranted to assess the association between alcohol consumption and the risk of developing cataracts.

1. Research Objectives:

The purpose of this study is to examine if alcohol, fruit and vegetable consumption have an effect on cataracts among adults aged 40 years or older in Canada.

The primary objectives are:

1. To examine if alcohol consumption affects the risk of cataracts.
2. To examine if fruit and vegetable consumption affects the risk of cataracts.

The secondary objectives are:

1. To examine if alcohol consumption and smoking interact to have an effect on the risk of cataracts.
2. To examine if alcohol consumption and gender interact to have an effect on the risk of cataracts.

2. Proposed data analysis and software requirements:

The first step in our analysis is to demonstrate descriptive statistics of cataracts and independent variables by using frequency tables (see table 1). The second objective will be addressed by using discrete-time survival analysis with time-varying lagged predictors. As the status of alcohol, fruit and vegetable consumption was updated at each cycle, the dataset would be lagged so that alcohol, fruit and vegetable at time t (cycles 1 through 8) predicted future incident cataract at time t+1 (cycle 2 through 9) to strengthen temporal arguments. The model will be adjusted by potential confounders, including age, gender, BMI, smoking, diabetes, hypertension, heart disease, socioeconomic status (education, income), physical activity, duration of oral

contraceptive, use of HT for menopause, steroid use, eye specialist visit, multivitamin use, UV exposure and glaucoma. The last step will examine the interaction between alcohol and smoking and gender by including an interaction term in our model. All data analysis will be carried out using SAS v 9.4 (SAS Institute Inc., Cary, NC, USA).

3. Data requirements:

I am requesting to use the National Population Health Survey (NPHS)- Household-Longitudinal component, cycle 1 through 9 confidential data files. This dataset includes self-reported cataracts and potential risk factors of cataracts. In addition, the longitudinal design of this survey is appropriate for my research to identify specific risk factors of cataracts. I will consult with the staff at the Western University RDC to ensure that acceptable sample sizes can be derived from each variable and that appropriate weighting and bootstrapping procedures are applied to the data.

4. Population of interest:

The population of interest in this study includes NPHS respondents aged 40 years and older in Canada.

5. Variables:

A number of variables have been identified from the 1996/1997 NPHS master file to meet the objectives of the study. A detailed list of variables will be found in Table1.

6. Expected project start and end dates:

This project is expected to begin in August 2019 and end in September 2020.

7. Expected products:

One or two peer-reviewed journal articles and a graduate-level thesis are expected as a result of this study.

8. Table 1: Requested variables

Variable Code	Variable Name
Cataracts(Depedent variable)	
CCC4_1S	Has cataracts

Alcohol drinking(Independent variable)	
All AL and AD regular All AL and AD derived	
Nutrition (Independent variable)	
All FV regular All FV derived	Fruit and Vegetable Consumption
NU	Supplement use
Socio-Economic (independent)	
DHC6_AGE	Age
SEX	Sex
INC4DHH	Derived total household income from all sources
DHC4_MAR	Marital status - DHC4_MAR
EDCnD1	Education
BMI(Independent)	
HWCnDBMI	BMI
Chronic conditions(Independent variable)	
CCC4_1F	Has high blood pressure
CCC4_1J	Has diabetes
CCC4_1L	Has heart disease
CCC4_1T	Has glaucoma
Physical activities(Independent variable)	
All PA and PA derived	Physical activities
Smoking(Independent variable)	
All SM and GH	Smoking and attitudes
UV exposure and Tanning(Independent variable)	

TU	UV exposure
Medication use(Independent variable)	
DGC6_1M	Took steroids
DGC6_1S	Took birth control pills
DGC6_1T	Took hormones for menopause

References

Bagnardi, V., Rota, M., Botteri, E., Tramacere, I., Islami, F., Fedirko, V., … La Vecchia, C. (2015). Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. British Journal of Cancer, 112(3), 580–593. https://doi.org/10.1038/bjc.2014.579

Bourne, R. R. A., Flaxman, S. R., Braithwaite, T., Cicinelli, M. V, Das, A., Jonas, J. B., … Zheng, Y. (2017). Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. The Lancet Global Health, 5(9), e888–e897. https://doi.org/https://doi.org/10.1016/S2214-109X(17)30293-0

Chasan-Taber, L., Willett, W. C., Seddon, J. M., Stampfer, M. J., Rosner, B., Colditz, G. A., … Hankinson, S. E. (2000). A prospective study of alcohol consumption and cataract extraction among U.S. women. Annals of Epidemiology, 10(6), 347–353.

Cumming, R. G., & Mitchell, P. (1997). Alcohol, smoking, and cataracts: the Blue Mountains Eye Study. Archives of Ophthalmology (Chicago, Ill. : 1960), 115(10), 1296–1303. https://doi.org/10.1001/archopht.1997.01100160466015

Echebiri, S. I., Odeigah, P. G. C., & Myers, S. N. (2010). Case-control studies and risk factors for cataract in two population studies in Nigeria. Middle East African Journal of Ophthalmology, 17(4), 303–309. https://doi.org/10.4103/0974-9233.71592

Floud, S., Kuper, H., Reeves, G. K., Beral, V., & Green, J. (2016). Risk Factors for Cataracts Treated Surgically in Postmenopausal Women. Ophthalmology, 123(8), 1704–1710.

https://doi.org/10.1016/j.ophtha.2016.04.037

Foster, A. (Ed.). (2000). Vision 2020: the cataract challenge. Community Eye Health, 13(34), 17–19. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17491949

Huang, G., Wu, L., Qiu, L., Lai, J., Huang, Z., & Liao, L. (2015). Association between vegetables consumption and the risk of age-related cataract: a meta-analysis. International Journal of Clinical and Experimental Medicine, 8(10), 18455–18461.

Kanthan, G. L., Mitchell, P., Burlutsky, G., & Wang, J. J. (2010). Alcohol consumption and the long-term incidence of cataract and cataract surgery: the Blue Mountains Eye Study.

American Journal of Ophthalmology, 150(3), 434-440.e1. https://doi.org/10.1016/j.ajo.2010.04.020

Klein, B. E. K., Klein, R., Lee, K. E., & Meuer, S. M. (2003). Socioeconomic and lifestyle factors and the 10-year incidence of age-related cataracts. American Journal of Ophthalmology, 136(3), 506–512. https://doi.org/10.1016/s0002-9394(03)00290-3

Klein, R., & Klein, B. E. K. (2013). The prevalence of age-related eye diseases and visual impairment in aging: current estimates. Investigative Ophthalmology & Visual Science, 54(14), ORSF5–ORSF13. https://doi.org/10.1167/iovs.13-12789

Kuang, T.-M., Tsai, S.-Y., Liu, C. J.-L., Ko, Y.-C., Lee, S.-M., & Chou, P. (2013). Seven-year incidence of age-related cataracts among an elderly Chinese population in Shihpai, Taiwan: The Shihpai Eye Study. Investigative Ophthalmology & Visual Science, 54(9), 6409–6415. https://doi.org/10.1167/iovs.13-12582

Li, L., Wan, X., & Zhao, G. (2014). Meta-analysis of the risk of cataract in type 2 diabetes. BMC Ophthalmology, 14, 94. https://doi.org/10.1186/1471-2415-14-94

Lindblad, B. E., Hakansson, N., Philipson, B., & Wolk, A. (2007). Alcohol consumption and risk of cataract extraction: a prospective cohort study of women. Ophthalmology, 114(4), 680–685. https://doi.org/10.1016/j.ophtha.2006.07.046

Liu, Y.-C., Wilkins, M., Kim, T., Malyugin, B., & Mehta, J. S. (2017). Cataracts. Lancet (London, England), 390(10094), 600–612. https://doi.org/10.1016/S0140-6736(17)30544-5

Manson, J. E., Christen, W. G., Seddon, J. M., Glynn, R. J., & Hennekens, C. H. (1994). A prospective study of alcohol consumption and risk of cataract. American Journal of Preventive Medicine, 10(3), 156–161.

McCarty, C. A., Mukesh, B. N., Fu, C. L., & Taylor, H. R. (1999). The epidemiology of cataract in Australia. American Journal of Ophthalmology, 128(4), 446–465. https://doi.org/10.1016/s0002-9394(99)00218-4

Morris, M. S., Jacques, P. F., Hankinson, S. E., Chylack, L. T. J., Willett, W. C., & Taylor, A. (2004). Moderate alcoholic beverage intake and early nuclear and cortical lens opacities. Ophthalmic Epidemiology, 11(1), 53–65. https://doi.org/10.1076/opep.11.1.53.26439

Munoz, B., Tajchman, U., Bochow, T., & West, S. (1993). Alcohol use and risk of posterior subcapsular opacities. Archives of Ophthalmology (Chicago, Ill. : 1960), 111(1), 110–112. https://doi.org/10.1001/archopht.1993.01090010114036

Pan, C.-W., Cheng, C.-Y., Saw, S.-M., Wang, J. J., & Wong, T. Y. (2013). Myopia and age-related cataract: a systematic review and meta-analysis. American Journal of Ophthalmology, 156(5), 1021-1033.e1. https://doi.org/10.1016/j.ajo.2013.06.005

Phillips, C. I., Clayton, R. M., Cuthbert, J., Qian, W., Donnelly, C. A., & Prescott, R. J. (1996).

Human cataract risk factors: significance of abstention from, and high consumption of, ethanol (U-curve) and non-significance of smoking. Ophthalmic Research, 28(4), 237–247. https://doi.org/10.1159/000267909

Poh, S., Mohamed Abdul, R. B. B., Lamoureux, E. L., Wong, T. Y., & Sabanayagam, C. (2016). Metabolic syndrome and eye diseases. Diabetes Research and Clinical Practice, 113, 86– 100. https://doi.org/10.1016/j.diabres.2016.01.016

Rehm, J., Gmel Sr, G. E., Gmel, G., Hasan, O. S. M., Imtiaz, S., Popova, S., … Shuper, P. A. (2017). The relationship between different dimensions of alcohol use and the burden of disease-an update. Addiction (Abingdon, England), 112(6), 968–1001. https://doi.org/10.1111/add.13757

Richter, G. M., Choudhury, F., Torres, M., Azen, S. P., & Varma, R. (2012). Risk factors for

incident cortical, nuclear, posterior subcapsular, and mixed lens opacities: the Los Angeles Latino eye study. Ophthalmology, 119(10), 2040–2047. https://doi.org/10.1016/j.ophtha.2012.05.001

Shahbazi, S., Studnicki, J., & Warner-Hillard, C. W. (2015). A Cross-Sectional Retrospective Analysis of the Racial and Geographic Variations in Cataract Surgery. PloS One, 10(11), e0142459. https://doi.org/10.1371/journal.pone.0142459

Singh, S., Pardhan, S., Kulothungan, V., Swaminathan, G., Ravichandran, J. S., Ganesan, S., … Raman, R. (2019). The prevalence and risk factors for cataract in rural and urban India.

Indian Journal of Ophthalmology, 67(4), 477–483. https://doi.org/10.4103/ijo.IJO_1127_17

Storey, P., Munoz, B., Friedman, D., & West, S. (2013). Racial differences in lens opacity incidence and progression: the Salisbury Eye Evaluation (SEE) study. Investigative Ophthalmology & Visual Science, 54(4), 3010–3018. https://doi.org/10.1167/iovs.12-11412

Tavani, A., Negri, E., & La Vecchia, C. (1996). Food and nutrient intake and risk of cataract.

Annals of Epidemiology, 6(1), 41–46.

Taylor, H. R., West, S. K., Rosenthal, F. S., Muñoz, B., Newland, H. S., Abbey, H., & Emmett, E. A. (1988). Effect of ultraviolet radiation on cataract formation. New England Journal of Medicine, 319(22), 1429-1433.

Tian, Y., Wu, J., Xu, G., Shen, L., Yang, S., Mandiwa, C., … Wang, Y. (2015). Parity and the risk of cataract: a cross-sectional analysis in the Dongfeng-Tongji cohort study. The British Journal of Ophthalmology, 99(12), 1650–1654. https://doi.org/10.1136/bjophthalmol-2015-

306653

Tsai, S.-Y., Hsu, W.-M., Cheng, C.-Y., Liu, J.-H., & Chou, P. (2003). Epidemiologic study of age-related cataracts among an elderly Chinese population in Shih-Pai, Taiwan.

Ophthalmology, 110(6), 1089–1095. https://doi.org/10.1016/S0161-6420(03)00243-4 Ughade, S. N., Zodpey, S. P., & Khanolkar, V. A. (1998). Risk factors for cataract: a case control study. Indian Journal of Ophthalmology, 46(4), 221–227.

Varma, R., & Torres, M. (2004). Prevalence of lens opacities in Latinos: the Los Angeles Latino Eye Study. Ophthalmology, 111(8), 1449–1456. https://doi.org/10.1016/j.ophtha.2004.01.024

Wang, W., & Zhang, X. (2014). Alcohol intake and the risk of age-related cataracts: a meta-analysis of prospective cohort studies. PloS One, 9(9), e107820. https://doi.org/10.1371/journal.pone.0107820

World Population Prospects: the 2019 Revision. (2019). new york. Retrieved from https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf

Appendix D. STROBE statement

	Item No	Recommendation	Chapter and section
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Chapter 1,Chapter 2
Objectives	3	State specific objectives, including any prespecified hypotheses	Chapter 3
Methods			
Study design	4	Present key elements of study design early in the paper	Chapter 4 Section 4.1
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Chapter 4 Section 4.2
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	Chapter 4 Section 4.3
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Chapter 4 Section 4.4
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Chapter 4 Section 4.4

Bias	9	Describe any efforts to address potential sources of bias	Chapter4 Section 4.1-4.4
Study size	10	Explain how the study size was arrived at	Chapter 4 Section 4.3
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Chapter4 Section 4.4
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Chapter 4 Section 4.5
		(*b*) Describe any methods used to examine subgroups and interactions	Chapter 4 Section 4.6.5
		(*c*) Explain how missing data were addressed	Chapter 4 Section 4.6.2, 4.6.5
		(*d*) If applicable, explain how loss to follow-up was addressed	Chapter 4 Section 4.6.1
		(*e*) Describe any sensitivity analyses	Chapter 4 Section 4.6.5
Results			
Participants	13	(a) Report numbers of individuals at each stage of study—e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Chapter 5 Section 5.1.1
		(b) Give reasons for non-participation at each stage	Chapter 4 Figure 3, 4
		(c) Consider use of a flow diagram	Chapter 4 Figure 3, 4
Descriptive data	14	(a) Give characteristics of study participants (e.g. demographic, clinical,	Chapter 5 Section 5.1.1

		social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Chapter 5 Section 5.1.1
		(c) Summarise follow-up time (e.g., average and total amount)	Chapter 5 Section 5.1, 5.2
Outcome data	15	Report numbers of outcome events or summary measures over time	Chapter 5 Section 5.2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Chapter 5 Section 5.3, 5.6
		(*b*) Report category boundaries when continuous variables were categorized	Chapter 5 Section 5.1, 5.3, 5.6
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	Chapter 5 Section 5.4, 5.5, 5.7, 5.8
Discussion			
Key results	18	Summarise key results with reference to study objectives	Chapter 6 Section 6.1
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Chapter 6 Section 6.3
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Chapter 6 Section 6.1
Generalisability	21	Discuss the generalisability (external validity) of the study results	N/A
Other information			

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	N/A

Abbreviations: N/A=Not applicable

Appendix E. Overall incidence rate and cumulative incidence at different follow-up time (1996-2011)

Follow-up Time	Incident Events	Cumulative Incidence	Incidence Rates (per 1000 person-years)
2 years	310	3.1%	
4 years	605	6.4%	
6 years	864	9.7%	
8 years	1,120	13.2%	
10 years	1,337	16.6%	
15 years	1,794	25.8%	
17 years	1,978	30.3%	19.2

Abbreviations: CI=Confidence interval

Appendix F. Overall incidence rate and cumulative incidence at different follow-up time (2004-2011)

Follow-up Time	Incident Events	Cumulative Incidence	Incidence Rates (per 1000 person-years)
2 years	303	4.1%	
4 years	554	7.9%	
6 years	769	11.6%	
8 years	1,001	15.9%	
9 years	1,019	16.3%	19.7

Abbreviations: CI=Confidence interval

Appendix G. Crude and adjusted Hazard Ratio for effect modification of age in the effect of alcohol intake on cataracts for NPHS respondents (1994-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.99 to 1.04)	0.38	1.00 (0.84 to 1.16)	0.99
Age (years)	1.09 (1.08 to 1.09)	<.001¶	1.31 (1.27 to 1.36)	<.0001¶
Alcohol intake*age			1.000 (0.998 to 1.002)	0.95
Sex (female)			1.39 (1.24 to 1.55)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.90 (0.76 to 1.06)	0.19
some post-secondary graduation			1.21 (1.06 to 1.38)	<.01¶
post-secondary graduation			0.97 (0.85 to 1.12)	0.70
Income				
lowest income			Reference	
lower middle income			0.99 (0.87 to 1.13)	0.91
upper middle income			1.06 (0.92 to 1.22)	0.43
highest income			1.07 (0.88 to 1.30)	0.52
Race				
White			Reference	
Asian			1.22 (0.88 to 1.68)	0.24
Black			0.83 (0.41 to 1.67)	0.60
Other			1.36 (0.61 to 3.05)	0.46
Smoking (yes)			1.08 (0.95 to 1.23)	0.25
BMI (kg/m^2)			0.96 (0.89 to 1.04)	0.33
BMI^2 (kg/m^2)			1.001 (0.999 to 1.002)	0.27
Physical activity				
active			Reference	
moderate			1.13 (0.96 to 1.32)	0.16

inactive	1.06 (0.92 to 1.22)	0.43
Use of vitamin (yes)	1.08 (0.97 to 1.19)	0.15
Hypertension (yes)	1.09 (0.97 to 1.23)	0.14
Diabetes (yes)	1.75 (1.46 to 2.11)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.82 (0.46 to 1.47)	0.51
cycle3 (1998-1999)	0.84 (0.48 to 1.49)	0.52
cycle4 (2000-2001)	0.58 (0.25 to 1.31)	0.19
cycle5 (2002-2003)	1.60 (0.83 to 3.07)	0.16
cycle6 (2004-2005)	0.25 (0.04 to 1.81)	0.17
cycle7 (2006-2007)	0.92 (0.23 to 3.75)	0.91
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix H. Crude and adjusted Hazard Ratio for effect modification of sex in the effect of alcohol intake on cataracts for NPHS respondents (1994-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.99 to 1.04)	0.38	0.98 (0.93 to 1.04)	0.51
Sex (female)	1.59 (1.45 to 1.74)	<.0001¶	1.36 (1.25 to 1.48)	<.0001¶
Alcohol intake*sex			1.03 (0.97 to 1.09)	0.34
Age (years)			1.31 (1.25 to 1.37)	<.0001¶
Age^2 (years)			0.999 (0.998 to 0.999)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.90 (0.76 to 1.06)	0.19
some post-secondary graduation			1.21 (1.06 to 1.38)	0.01¶
post-secondary graduation			0.97 (0.85 to 1.12)	0.69
Income				
lowest income			Reference	
lower middle income			0.99 (0.87 to 1.13)	0.90
upper middle income			1.06 (0.92 to 1.22)	0.45
highest income			1.07 (0.88 to 1.30)	0.52
Race				
White			Reference	
Asian			1.21 (0.88 to 1.68)	0.24
Black			0.83 (0.41 to 1.67)	0.60
Other			1.36 (0.61 to 3.04)	0.46
Smoking (yes)			1.08 (0.95 to 1.22)	0.26
BMI (kg/m^2)			0.96 (0.89 to 1.04)	0.33
BMI^2 (kg/m^2)			1.001 (0.999 to 1.002)	0.27
Physical activity				
active			Reference	

moderate	1.12 (0.96 to 1.32)	0.16
inactive	1.06 (0.92 to 1.22)	0.44
Use of vitamin (yes)	1.08 (0.97 to 1.19)	0.15
Hypertension (yes)	1.09 (0.97 to 1.23)	0.16
Diabetes (yes)	1.76 (1.46 to 2.12)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.82 (0.46 to 1.47)	0.51
cycle3 (1998-1999)	0.84 (0.48 to 1.49)	0.54
cycle4 (2000-2001)	0.57 (0.25 to 1.31)	0.19
cycle5 (2002-2003)	1.60 (0.83 to 3.07)	0.16
cycle6 (2004-2005)	0.25 (0.04 to 1.80)	0.17
cycle7 (2006-2007)	0.92 (0.23 to 3.75)	0.91
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix I. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for NPHS respondents older than 65 (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.00 (0.97 to 1.03)	0.79	1.03 (0.95 to 1.11)	0.47
Alcohol intake^2 (10 drinks/month)			0.998 (0.992 to 1.004)	0.49
Sex (female)			1.22 (1.05 to 1.43)	0.01¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.84 (0.67 to 1.06)	0.15
some post-secondary graduation			1.10 (0.91 to 1.33)	0.34
post-secondary graduation			0.89 (0.73 to 1.10)	0.28
Income				
lowest income			Reference	
lower middle income			0.91 (0.77 to 1.07)	0.25
upper middle income			0.95 (0.78 to 1.17)	0.65
highest income			0.94 (0.66 to 1.33)	0.71
Race				
White			Reference	
Asian			0.87 (0.47 to 1.64)	0.67
Black			0.59 (0.19 to 1.85)	0.37
Other			1.07 (0.15 to 7.74)	0.94
Smoking (yes)			0.75 (0.61 to 0.94)	0.01¶
BMI (kg/m^2)			0.88 (0.79 to 0.97)	0.01¶
BMI^2 (kg/m^2)			1.002 (1.000 to 1.004)	0.02¶
Physical activity				
active			Reference	
moderate			1.18 (0.94 to 1.49)	0.16
inactive			1.15 (0.94 to 1.42)	0.17

Use of vitamin (yes)	1.21 (1.05 to 1.40)	0.01¶
Hypertension (yes)	1.05 (0.90 to 1.22)	0.53
Diabetes (yes)	1.29 (1.00 to 1.66)	0.05

Abbreviations: CI=Confidence interval, BMI=Body mass index

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix J. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for NPHS respondents equal or younger than 65 (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.02 (0.99 to 1.06)	0.18	1.01 (0.95 to 1.08)	0.80
Alcohol intake^2 (10 drinks/month)			1.00 (1.00 to 1.01)	0.29
Sex (female)			1.68 (1.44 to 1.96)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.78 (0.61 to 0.98)	0.04¶
some post-secondary graduation			1.01 (0.84 to 1.23)	0.88
post-secondary graduation			0.75 (0.62 to 0.91)	<.01¶
Income				
lowest income			Reference	
lower middle income			0.99 (0.80 to 1.23)	0.96
upper middle income			0.88 (0.71 to 1.09)	0.24
highest income			0.90 (0.70 to 1.16)	0.42
Race				
White			Reference	
Asian			1.24 (0.85 to 1.83)	0.27
Black			0.74 (0.31 to 1.78)	0.50
Other			0.98 (0.40 to 2.37)	0.96
Smoking (yes)			0.98 (0.84 to 1.15)	0.80
BMI (kg/m^2)			1.22 (1.07 to 1.38)	<.01¶
BMI^2 (kg/m^2)			1.00 (1.00 to 1.00)	<.01¶
Physical activity				
active			Reference	
moderate			1.02 (0.81 to 1.28)	0.90
inactive			0.89 (0.73 to 1.09)	0.27
Use of vitamin (yes)			0.99 (0.86 to 1.15)	0.92

Hypertension (yes)	1.60 (1.32 to 1.93)	<.0001¶
Diabetes (yes)	3.19 (2.42 to 4.21)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.21 (0.12 to 0.36)	<.0001¶
cycle3 (1998-1999)	0.22 (0.13 to 0.39)	<.0001¶
cycle4 (2000-2001)	0.16 (0.07 to 0.36)	<.0001¶
cycle5 (2002-2003)	0.47 (0.25 to 0.88)	0.02¶
cycle6 (2004-2005)	0.08 (0.01 to 0.56)	0.01¶
cycle7 (2006-2007)	0.34 (0.08 to 1.36)	0.13
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix K. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for female NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.03 (0.98 to 1.09)	0.19	1.10 (0.98 to 1.23)	0.11
Alcohol intake^2 (10 drinks/month)			0.99 (0.98 to 1.01)	0.23
Age (years)			1.31 (1.24 to 1.39)	<.0001¶
Age^2 (years)			0.998 (0.998 to 0.999)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.86 (0.71 to 1.05)	0.15
some post-secondary graduation			1.17 (1.00 to 1.38)	0.05
post-secondary graduation			0.96 (0.81 to 1.15)	0.67
Income				
lowest income			Reference	
lower middle income			1.01 (0.87 to 1.18)	0.88
upper middle income			1.04 (0.87 to 1.24)	0.66
highest income			0.94 (0.73 to 1.22)	0.65
Race				
White			Reference	
Asian			0.73 (0.43 to 1.21)	0.22
Black			0.60 (0.15 to 2.42)	0.47
Other			1.44 (0.54 to 3.87)	0.47

Smoking (yes)	1.09 (0.93 to 1.28)	0.29
BMI (kg/m^2)	0.96 (0.88 to 1.04)	0.33
BMI^2 (kg/m^2)	1.001 (0.999 to 1.002)	0.25
Physical activity		
active	Reference	
moderate	1.19 (0.96 to 1.48)	0.12
inactive	1.18 (0.97 to 1.43)	0.09
Use of vitamin (yes)	1.06 (0.94 to 1.20)	0.34
Hypertension (yes)	1.07 (0.93 to 1.24)	0.34
Diabetes (yes)	1.69 (1.31 to 2.18)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.57 (0.25 to 1.32)	0.19
cycle3 (1998-1999)	0.59 (0.25 to 1.34)	0.16
cycle4 (2000-2001)	0.75 (0.30 to 1.87)	0.54
cycle5 (2002-2003)	1.43 (0.62 to 3.32)	0.40
cycle6 (2004-2005)	0.38 (0.05 to 2.72)	0.33
cycle7 (2006-2007)	1.47 (0.36 to 6.04)	0.59
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix L. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts for male NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.04 (1.01 to 1.06)	0.01¶	0.97 (0.91 to 1.03)	0.27
Alcohol intake^2 (10 drinks/month)			1.003 (1.000 to 1.006)	0.06
Age (years)			1.29 (1.19 to 1.40)	<.0001¶
Age^2 (years)			1.00 (1.00 to 1.00)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			1.00 (0.74 to 1.34)	0.99
some post-secondary graduation			1.23 (0.97 to 1.56)	0.08
post-secondary graduation			0.96 (0.76 to 1.22)	0.75
Income				
lowest income			Reference	
lower middle income			0.90 (0.70 to 1.16)	0.42
upper middle income			1.05 (0.81 to 1.37)	0.69
highest income			1.24 (0.90 to 1.72)	0.19
Race				
White			Reference	
Asian			2.26 (1.47 to 3.49)	<.0001¶
Black			0.98 (0.44 to 2.20)	0.96
Other			1.10 (0.27 to 4.50)	0.89
Smoking (yes)			1.05 (0.85 to 1.30)	0.67
BMI (kg/m^2)			1.01 (0.82 to 1.23)	0.96
BMI^2 (kg/m^2)			1.00 (1.00 to 1.00)	0.95
Physical activity				
active			Reference	
moderate			1.10 (0.86 to 1.42)	0.45

inactive	0.93 (0.74 to 1.15)	0.49
Use of vitamin (yes)	1.10 (0.91 to 1.32)	0.33
Hypertension (yes)	1.13 (0.91 to 1.40)	0.28
Diabetes (yes)	1.82 (1.38 to 2.41)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	1.29 (0.57 to 2.91)	0.54
cycle3 (1998-1999)	1.30 (0.55 to 2.94)	0.48
cycle4 (2000-2001)	0.25 (0.03 to 1.83)	0.17
cycle5 (2002-2003)	1.89 (0.66 to 5.35)	0.23
cycle6 (2004-2005)	N/A	N/A
cycle7 (2006-2007)	N/A	N/A
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix M. Crude and adjusted Hazard Ratio for the association between alcohol intake and cataracts with competing risk for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Alcohol intake (10 drinks/month)	1.01 (0.98 to 1.03)	0.71	1.00 (0.94 to 1.06)	0.87
Alcohol intake^2 (10 drinks/month)			1.001 (0.997 to 1.005)	0.69
Age (years)			1.33 (1.27 to 1.40)	<.0001¶
Age^2 (years)			0.998 (0.998 to 0.999)	<.0001¶
Sex (female)			1.38 (1.24 to 1.54)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.92 (0.78 to 1.08)	0.29
some post-secondary graduation			1.21 (1.06 to 1.37)	0.01¶
post-secondary graduation			0.98 (0.85 to 1.13)	0.79
Income				
lowest income			Reference	
lower middle income			1.02 (0.89 to 1.17)	0.76
upper middle income			1.08 (0.94 to 1.25)	0.30
highest income			1.09 (0.90 to 1.32)	0.38
Race				
White			Reference	
Asian			1.23 (0.87 to 1.73)	0.24
Black			0.85 (0.42 to 1.73)	0.66
Other			1.37 (0.60 to 3.11)	0.45
Smoking (yes)			1.06 (0.94 to 1.21)	0.34
BMI (kg/m^2)			0.96 (0.89 to 1.03)	0.21
BMI^2 (kg/m^2)			1.001 (1.000 to 1.002)	0.16
Physical activity				
active			Reference	
moderate			1.12 (0.95 to 1.31)	0.17

inactive	1.05 (0.91 to 1.21)	0.53
Use of vitamin (yes)	1.07 (0.96 to 1.18)	0.22
Hypertension (yes)	1.09 (0.97 to 1.23)	0.16
Diabetes (yes)	1.71 (1.40 to 2.08)	<.0001¶
Year of cohort entry		
cycle1 (1994-1995)	Reference	
cycle2 (1996-1997)	0.85 (0.48 to 1.51)	0.57
cycle3 (1998-1999)	0.86 (0.49 to 1.53)	0.58
cycle4 (2000-2001)	0.59 (0.26 to 1.33)	0.20
cycle5 (2002-2003)	1.61 (0.84 to 3.08)	0.15
cycle6 (2004-2005)	0.25 (0.04 to 1.82)	0.17
cycle7 (2006-2007)	0.92 (0.23 to 3.72)	0.91
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix N. Crude and adjusted Hazard Ratio for age difference in the effect of fruit and vegetable consumption on cataracts for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.06 (1.04 to 1.09)	<.0001¶	0.97 (0.68 to 1.27)	0.85
Age (years)	1.07 (1.07 to 1.08)	<.0001¶	1.48 (1.40 to 1.57)	<.0001¶
Fruit and vegetable consumption*Age			1.001 (0.997 to 1.005)	0.60
Alcohol intake (10 drinks/month)			0.92 (0.63 to 1.32)	0.64
Alcohol intake^2 (10 drinks/month)			1.01 (0.96 to 1.06)	0.76
Sex (female)			1.60 (1.32 to 1.93)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.96 (0.71 to 1.28)	0.76
some post-secondary graduation			1.20 (0.94 to 1.51)	0.14
post-secondary graduation			1.04 (0.82 to 1.32)	0.73
Income				
lowest income			Reference	
lower middle income			0.96 (0.72 to 1.30)	0.81
upper middle income			1.00 (0.74 to 1.34)	0.98
highest income			0.87 (0.62 to 1.23)	0.44
Race				
White			Reference	
Asian			0.92 (0.45 to 1.86)	0.82
Black			1.07 (0.27 to 4.33)	0.92
Other			1.83 (0.58 to 5.77)	0.30
Smoking (yes)			1.13 (0.89 to 1.43)	0.32
BMI (kg/m^2)			1.03 (0.89 to 1.20)	0.66
BMI^2 (kg/m^2)			1.000 (0.997 to 1.002)	0.76
Physical activity				

active	Reference	
moderate	1.13 (0.89 to 1.45)	0.32
inactive	0.96 (0.76 to 1.22)	0.75
Use of vitamin (yes)	1.07 (0.90 to 1.28)	0.42
Hypertension (yes)	1.02 (0.85 to 1.24)	0.82
Diabetes (yes)	1.15 (0.86 to 1.55)	0.34
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.05 (0.25 to 4.44)	0.95
cycle7 (2006-2007)	2.66 (0.80 to 8.84)	0.11
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix O. Crude and adjusted Hazard Ratio for effect modification of sex in the effect of fruit and vegetable consumption on cataracts for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.06 (1.04 to 1.09)	<.0001¶	1.06 (0.92 to 1.20)	0.41
Sex (female)	1.63 (1.43 to 1.85)	<.0001¶	1.85 (1.40 to 2.30)	0.01¶
Fruit and vegetable consumption*Sex (female)			0.97 (0.87 to 1.06)	0.48
Alcohol intake (10 drinks/month)			0.92 (0.64 to 1.34)	0.67
Alcohol intake^2 (10 drinks/month)			1.01 (0.96 to 1.06)	0.78
Age (years)			1.49 (1.37 to 1.62)	<.0001¶
Age^2 (years)			0.998 (0.997 to 0.998)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.95 (0.71 to 1.28)	0.75
some post-secondary graduation			1.20 (0.95 to 1.52)	0.13
post-secondary graduation			1.04 (0.82 to 1.33)	0.72
Income				
lowest income			Reference	
lower middle income			0.97 (0.72 to 1.30)	0.81
upper middle income			1.00 (0.75 to 1.34)	1.00
highest income			0.88 (0.63 to 1.23)	0.45
Race				
White			Reference	
Asian			0.93 (0.46 to 1.87)	0.83
Black			1.07 (0.27 to 4.34)	0.92
Other			1.84 (0.58 to 5.78)	0.30
Smoking (yes)			1.13 (0.89 to 1.42)	0.33
BMI (kg/m^2)			1.04 (0.89 to 1.20)	0.65

BMI^2 (kg/m^2)	1.000 (0.997 to 1.002)	0.74
Physical activity		
active	Reference	
moderate	1.14 (0.89 to 1.46)	0.30
inactive	0.97 (0.76 to 1.22)	0.78
Use of vitamin (yes)	1.08 (0.90 to 1.28)	0.42
Hypertension (yes)	1.03 (0.85 to 1.24)	0.80
Diabetes (yes)	1.16 (0.86 to 1.55)	0.33
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.04 (0.25 to 4.39)	0.96
cycle7 (2006-2007)	2.63 (0.79 to 8.74)	0.11
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix P. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts for NPHS respondents older than 65 (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.01 (0.97 to 1.05)	0.61	1.00 (0.87 to 1.16)	0.97
Fruit and vegetable consumption^2 (servings/day)			1.00 (0.99 to 1.01)	0.90
Alcohol intake (10 drinks/month)			1.11 (0.64 to 1.94)	0.70
Alcohol intake^2 (10 drinks/month)			0.99 (0.92 to 1.07)	0.87
Sex (female)			1.51 (1.17 to 1.96)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			1.00 (0.69 to 1.45)	0.99
some post-secondary graduation			1.16 (0.85 to 1.57)	0.35
post-secondary graduation			1.01 (0.73 to 1.39)	0.96
Income				
lowest income			Reference	
lower middle income			1.02 (0.72 to 1.45)	0.90
upper middle income			1.02 (0.71 to 1.47)	0.93
highest income			0.84 (0.52 to 1.34)	0.46
Race				
White			Reference	
Asian			0.44 (0.11 to 1.80)	0.26
Black			N/A	N/A
Other			1.94 (0.26 to 14.39)	0.52
Smoking (yes)			1.00 (0.71 to 1.41)	0.99
BMI (kg/m^2)			1.20 (0.92 to 1.55)	0.18
BMI^2 (kg/m^2)			1.00 (0.99 to 1.00)	0.25
Physical activity				

active	Reference	
moderate	1.21 (0.86 to 1.70)	0.29
inactive	1.01 (0.73 to 1.41)	0.93
Use of vitamin (yes)	0.93 (0.74 to 1.18)	0.56
Hypertension (yes)	1.01 (0.80 to 1.29)	0.91
Diabetes (yes)	0.99 (0.68 to 1.45)	0.97

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix Q. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts for NPHS respondents equal or younger than 65 (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.07 (1.02 to 1.12)	0.01¶	1.30 (0.99 to 1.71)	0.06
Fruit and vegetable consumption^2 (servings/day)			0.98 (0.95 to 1.00)	0.09
Alcohol intake (10 drinks/month)			0.69 (0.42 to 1.14)	0.15
Alcohol intake^2 (10 drinks/month)			1.04 (0.97 to 1.11)	0.33
Sex (female)			1.63 (1.22 to 2.16)	<.01¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.77 (0.47 to 1.26)	0.29
some post-secondary graduation			0.96 (0.65 to 1.42)	0.82
post-secondary graduation			0.82 (0.56 to 1.20)	0.31
Income				
lowest income			Reference	
lower middle income			0.79 (0.45 to 1.38)	0.41
upper middle income			0.77 (0.46 to 1.29)	0.31
highest income			0.63 (0.37 to 1.08)	0.09
Race				
White			Reference	
Asian			1.09 (0.48 to 2.47)	0.84
Black			1.05 (0.26 to 4.26)	0.95
Other			1.61 (0.39 to 6.55)	0.51
Smoking (yes)			0.99 (0.72 to 1.36)	0.94
BMI (kg/m^2)			1.02 (0.84 to 1.25)	0.81
BMI^2 (kg/m^2)			1.00 (1.00 to 1.00)	0.69
Physical activity				

active	Reference	
moderate	1.09 (0.76 to 1.57)	0.63
inactive	1.02 (0.73 to 1.44)	0.90
Use of vitamin (yes)	1.37 (1.06 to 1.78)	0.02¶
Hypertension (yes)	1.80 (1.32 to 2.46)	<.01¶
Diabetes (yes)	2.32 (1.47 to 3.68)	<.01¶
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	0.17 (0.04 to 0.69)	0.01¶
cycle7 (2006-2007)	0.54 (0.17 to 1.70)	0.29
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix R. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts for female NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.04 (1.00 to 1.07)	0.04¶	1.04 (0.88 to 1.23)	0.66
Fruit and vegetable consumption^2 (servings/day)			1.00 (0.98 to 1.01)	0.53
Alcohol intake (10 drinks/month)			0.68 (0.35 to 1.30)	0.24
Alcohol intake^2 (10 drinks/month)			1.12 (0.97 to 1.30)	0.13
Age (years)			1.49 (1.34 to 1.65)	<.0001¶
Age^2 (years)			0.998 (0.997 to 0.998)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.86 (0.60 to 1.25)	0.43
some post-secondary graduation			1.18 (0.88 to 1.58)	0.26
post-secondary graduation			1.00 (0.74 to 1.36)	0.99
Income				
lowest income			Reference	
lower middle income			1.21 (0.85 to 1.72)	0.29
upper middle income			1.11 (0.77 to 1.60)	0.59
highest income			0.94 (0.61 to 1.45)	0.79
Race				
White			Reference	
Asian			0.17 (0.02 to 1.19)	0.07
Black			2.18 (0.30 to 15.72)	0.44
Other			1.56 (0.38 to 6.36)	0.54
Smoking (yes)			1.07 (0.79 to 1.44)	0.67
BMI (kg/m^2)			0.97 (0.83 to 1.13)	0.67
BMI^2 (kg/m^2)			1.001 (0.998 to 1.003)	0.62

Physical activity		
active	Reference	
moderate	1.09 (0.79 to 1.50)	0.60
inactive	0.98 (0.73 to 1.33)	0.92
Use of vitamin (yes)	1.05 (0.85 to 1.31)	0.63
Hypertension (yes)	0.97 (0.77 to 1.23)	0.82
Diabetes (yes)	1.04 (0.69 to 1.56)	0.85
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.42 (0.33 to 6.19)	0.64
cycle7 (2006-2007)	3.63 (1.06 to 12.50)	0.04
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix S. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts for male NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.07 (1.01 to 1.13)	0.02¶	1.01 (0.79 to 1.30)	0.91
Fruit and vegetable consumption^2 (servings/day)			1.00 (0.98 to 1.02)	0.89
Alcohol intake (10 drinks/month)			0.98 (0.58 to 1.66)	0.94
Alcohol intake^2 (10 drinks/month)			1.00 (0.93 to 1.07)	1.00
Age (years)			1.49 (1.29 to 1.72)	<.0001¶
Age^2 (years)			1.00 (1.00 to 1.00)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			1.14 (0.70 to 1.87)	0.60
some post-secondary graduation			1.21 (0.80 to 1.81)	0.37
post-secondary graduation			1.11 (0.75 to 1.64)	0.61
Income				
lowest income			Reference	
lower middle income			0.54 (0.31 to 0.94)	0.03¶
upper middle income			0.78 (0.47 to 1.28)	0.32
highest income			0.72 (0.41 to 1.26)	0.25
Race				
White			Reference	
Asian			2.68 (1.20 to 5.96)	0.02
Black			0.74 (0.10 to 5.34)	0.76
Other			4.17 (0.56 to 31.19)	0.16
Smoking (yes)			1.25 (0.85 to 1.83)	0.26
BMI (kg/m^2)			1.71 (1.06 to 2.75)	0.03
BMI^2 (kg/m^2)			0.99 (0.98 to 1.00)	0.04¶

Physical activity		
active	Reference	
moderate	1.23 (0.83 to 1.81)	0.31
inactive	0.92 (0.63 to 1.34)	0.65
Use of vitamin (yes)	1.11 (0.83 to 1.50)	0.49
Hypertension (yes)	1.07 (0.78 to 1.48)	0.66
Diabetes (yes)	1.34 (0.87 to 2.05)	0.19

Abbreviations: CI=Confidence interval, BMI=Body mass index, NA=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix T. Crude and adjusted Hazard Ratio for the association between fruit and vegetable consumption and cataracts with competing risk for NPHS respondents (2002-2011)

Variable	Crude Hazard Ratio (95% CI)	P-value	Adjusted Hazard Ratio (95% CI)*	P-value
Fruit and vegetable consumption (servings/day)	1.06 (1.04 to 1.09)	<.0001¶	1.04 (0.90 to 1.19)	0.63
Fruit and vegetable consumption^2 (servings/day)			1.00 (0.99 to 1.01)	0.57
Alcohol intake (10 drinks/month)			0.93 (0.64 to 1.34)	0.69
Alcohol intake^2 (10 drinks/month)			1.01 (0.97 to 1.05)	0.73
Age (years)			1.49 (1.36 to 1.62)	<.0001¶
Age^2 (years)			0.998 (0.997 to 0.998)	<.0001¶
Sex (female)			1.60 (1.33 to 1.93)	<.0001¶
Education				
less than secondary graduation			Reference	
secondary graduation			0.97 (0.73 to 1.29)	0.83
some post-secondary graduation			1.20 (0.96 to 1.50)	0.11
post-secondary graduation			1.06 (0.84 to 1.33)	0.64
Income				
lowest income			Reference	
lower middle income			0.99 (0.74 to 1.33)	0.95
upper middle income			1.02 (0.76 to 1.38)	0.88
highest income			0.90 (0.64 to 1.27)	0.55
Race				
White			Reference	
Asian			0.93 (0.44 to 1.99)	0.86
Black			1.07 (0.28 to 4.08)	0.92
Other			1.82 (0.64 to 5.17)	0.26
Smoking (yes)			1.10 (0.87 to 1.38)	0.42
BMI (kg/m^2)			1.04 (0.90 to 1.19)	0.63

BMI^2 (kg/m^2)	1.000 (0.997 to 1.002)	0.73
Physical activity		
active	Reference	
moderate	1.13 (0.89 to 1.44)	0.31
inactive	0.96 (0.77 to 1.21)	0.75
Use of vitamin (yes)	1.08 (0.91 to 1.28)	0.39
Hypertension (yes)	1.03 (0.85 to 1.23)	0.79
Diabetes (yes)	1.16 (0.87 to 1.55)	0.31
Year of cohort entry		
cycle5 (2002-2003)	Reference	
cycle6 (2004-2005)	1.03 (0.25 to 4.35)	0.97
cycle7 (2006-2007)	2.62 (0.79 to 8.73)	0.12
cycle8 (2008-2009)	N/A	N/A

Abbreviations: CI=Confidence interval, BMI=Body mass index, N/A=Not applicable

¶ Indicates a P-value<0.05

*Adjusted for FV consumption, alcohol intake, age, sex, education, income, race, smoking, BMI, physical activity, vitamin use, hypertension, diabetes and year of cohort entry

Appendix U. Missing data pattern for cohort 1

Group	Education	Smoking	Use of vitamin	Hypertension	Diabetes	Physical activity	Sex	BMI	Income	Age	Alcohol intake	Race	Year of cohort entry	Percent
1	X	X	X	X	X	X	X	X	X	X	X	X	X	75.11%
2	X	X	X	X	X	X	X	X	X	X	X	.	X	6.92%
3	X	X	X	X	X	X	X	X	X	X	.	X	X	1.79%
4	X	X	X	X	X	X	X	X	X	X	.	.	X	0.08%
5	X	X	X	X	X	X	X	X	.	X	X	X	X	3.54%
6	X	X	X	X	X	X	X	X	.	X	X	.	X	0.34%
7	X	X	X	X	X	X	X	X	.	X	.	X	X	0.11%
8	X	X	X	X	X	X	X	.	X	X	X	X	X	0.77%
9	X	X	X	X	X	X	X	.	X	X	X	.	X	0.03%
10	X	X	X	X	X	X	X	.	X	X	.	X	X	0.04%
11	X	X	X	X	X	X	X	.	X	X	.	.	X	0.01%
12	X	X	X	X	X	X	X	.	.	X	X	X	X	0.02%
13	X	X	X	X	X	X	X	.	.	X	.	X	X	0.01%
14	X	X	X	X	X	.	X	X	X	X	X	X	X	2.41%
15	X	X	X	X	X	.	X	X	X	X	X	.	X	0.33%
16	X	X	X	X	X	.	X	X	X	X	.	X	X	0.03%
17	X	X	X	X	X	.	X	X	X	X	.	.	X	0.01%
18	X	X	X	X	X	.	X	X	.	X	X	X	X	0.23%
19	X	X	X	X	X	.	X	X	.	X	X	.	X	0.02%
20	X	X	X	X	X	.	X	.	X	X	X	X	X	0.02%
21	X	X	X	X	X	.	X	.	X	X	.	X	X	0.03%
22	X	X	X	.	X	X	X	X	X	X	X	X	X	0.01%
23	X	X	X	.	X	X	X	X	X	X	X	.	X	0.01%
24	X	X	.	X	X	X	X	X	X	X	X	X	X	0.21%
25	X	X	.	X	X	X	X	X	X	X	X	.	X	0.04%
26	X	X	.	X	X	X	X	X	X	X	.	X	X	0.01%
27	X	X	.	X	X	X	X	X	.	X	X	X	X	0.04%
28	X	X	.	X	X	X	X	.	X	X	X	X	X	0.01%
29	X	X	.	X	X	.	X	X	X	X	X	X	X	0.02%
30	X	X	.	X	X	.	X	X	.	X	X	X	X	0.02%
31	X	X	.	X	X	.	X	X	.	X	X	.	X	0.01%
32	X	X	.	X	X	.	X	.	X	X	X	X	X	0.01%

33	X	.	X	X	X	X	X	X	X	X	X	X	X	4.78%
34	X	.	X	X	X	X	X	X	X	X	X	.	X	0.06%
35	X	.	X	X	X	X	X	X	X	X	.	X	X	0.11%
36	X	.	X	X	X	X	X	X	.	X	X	X	X	0.2%
37	X	.	X	X	X	X	X	X	.	X	.	X	X	0.04%
38	X	.	X	X	X	X	X	.	X	X	X	X	X	0.01%
39	X	.	X	X	X	.	X	X	X	X	X	X	X	0.03%
40	X	.	X	X	X	.	X	X	.	X	X	.	X	0.01%
41	X	.	X	X	X	.	X	.	X	X	X	X	X	0.03%
42	X	.	.	X	X	.	X	X	X	X	X	X	X	0.03%
43	X	.	.	X	X	.	X	X	X	X	X	.	X	0.01%
44	X	.	.	X	X	.	X	X	X	X	.	X	X	0.79%
45	X	.	.	X	X	.	X	X	X	X	.	.	X	0.1%
46	X	.	.	X	X	.	X	X	.	X	X	X	X	0.02%
47	X	.	.	X	X	.	X	X	.	X	.	X	X	0.16%
48	X	.	.	X	X	.	X	X	.	X	.	.	X	0.06%
49	X	.	.	X	X	.	X	.	X	X	.	X	X	0.54%
50	X	.	.	X	X	.	X	.	X	X	.	.	X	0.11%
51	X	.	.	X	X	.	X	.	.	X	X	X	X	0.01%
52	X	.	.	X	X	.	X	.	.	X	.	X	X	0.06%
53	X	.	.	X	X	.	X	.	.	X	.	.	X	0.02%
54	.	X	X	X	X	X	X	X	X	X	X	X	X	0.06%
55	.	X	X	X	X	X	X	X	X	X	X	.	X	0.03%
56	.	X	X	X	X	X	X	X	.	X	X	X	X	0.09%
57	.	X	X	X	X	X	X	.	.	X	.	X	X	0.01%
58	.	X	X	X	X	.	X	X	X	X	X	X	X	0.01%
59	.	X	X	X	X	.	X	X	.	X	X	X	X	0.03%
60	.	X	.	X	X	X	X	X	.	X	X	X	X	0.01%
61	.	X	.	X	X	X	X	X	.	X	X	.	X	0.01%
62	.	X	.	X	X	X	X	X	.	X	.	X	X	0.01%
63	.	X	.	X	X	.	X	X	.	X	X	X	X	0.03%
64	.	X	.	X	X	.	X	X	.	X	.	X	X	0.01%
65	.	.	.	X	X	X	X	X	X	X	.	X	X	0.01%
66	.	.	.	X	X	X	X	X	.	X	X	X	X	0.07%
67	.	.	.	X	X	X	X	X	.	X	.	.	X	0.01%
68	.	.	.	X	X	.	X	X	.	X	X	X	X	0.1%
69	.	.	.	X	X	.	X	X	.	X	X	.	X	0.01%

70	.	.	.	X	X	.	X	X	.	X	.	X	X	0.04%
71	.	.	.	X	X	.	X	.	X	X	.	X	X	0.01%

Appendix V. Missing data pattern for cohort 2

Group	Education	Smoking	Use of vitamin	Hypertension	Diabetes	Physical activity	Sex	BMI	Income	Age	Fruit and vegetable consumption	Alcohol intake	Race	Year of cohort entry	Percent
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	64.75%
2	X	X	X	X	X	X	X	X	X	X	X	X	.	X	1.1%
3	X	X	X	X	X	X	X	X	X	X	X	.	X	X	17.31%
4	X	X	X	X	X	X	X	X	X	X	X	.	.	X	0.37%
5	X	X	X	X	X	X	X	X	X	X	.	X	X	X	0.41%
6	X	X	X	X	X	X	X	X	X	X	.	X	.	X	0.04%
7	X	X	X	X	X	X	X	X	X	X	.	.	X	X	0.24%
8	X	X	X	X	X	X	X	X	X	X	.	.	.	X	0.01%
9	X	X	X	X	X	X	X	X	.	X	X	X	X	X	4.52%
10	X	X	X	X	X	X	X	X	.	X	X	X	.	X	0.14%
11	X	X	X	X	X	X	X	X	.	X	X	.	X	X	2.41%
12	X	X	X	X	X	X	X	X	.	X	X	.	.	X	0.07%
13	X	X	X	X	X	X	X	X	.	X	.	X	X	X	0.11%
14	X	X	X	X	X	X	X	X	.	X	.	.	X	X	0.12%
15	X	X	X	X	X	X	X	.	X	X	X	X	X	X	1.89%
16	X	X	X	X	X	X	X	.	X	X	X	X	.	X	0.05%
17	X	X	X	X	X	X	X	.	X	X	X	.	X	X	0.57%
18	X	X	X	X	X	X	X	.	X	X	.	.	X	X	0.01%
19	X	X	X	X	X	X	X	.	.	X	X	X	X	X	0.11%
20	X	X	X	X	X	X	X	.	.	X	X	.	X	X	0.07%
21	X	X	X	X	X	X	X	.	.	X	.	.	X	X	0.01%
22	X	X	X	X	X	.	X	X	X	X	X	.	X	X	0.01%
23	X	X	X	X	X	.	X	X	X	X	.	X	X	X	0.99%
24	X	X	X	X	X	.	X	X	X	X	.	X	.	X	0.03%
25	X	X	X	X	X	.	X	X	X	X	.	.	X	X	0.88%
26	X	X	X	X	X	.	X	X	X	X	.	.	.	X	0.05%
27	X	X	X	X	X	.	X	X	.	X	X	X	.	X	0.01%
28	X	X	X	X	X	.	X	X	.	X	X	.	X	X	0.01%
29	X	X	X	X	X	.	X	X	.	X	.	X	X	X	0.19%
30	X	X	X	X	X	.	X	X	.	X	.	.	X	X	0.22%
31	X	X	X	X	X	.	X	X	.	X	.	.	.	X	0.01%
32	X	X	X	X	X	.	X	.	X	X	.	X	X	X	0.07%

33	X	X	X	X	X	.	X	.	X	X	.	.	X	X	0.11%
34	X	X	X	X	.	X	X	.	X	X	X	.	X	X	0.01%
35	X	X	X	.	X	X	X	X	X	X	X	X	X	X	0.07%
36	X	X	X	.	X	X	X	X	X	X	X	.	X	X	0.03%
37	X	X	X	.	X	X	X	X	.	X	X	X	X	X	0.01%
38	X	X	X	.	X	X	X	X	.	X	.	.	X	X	0.01%
39	X	X	X	.	X	.	X	X	.	X	.	X	X	X	0.01%
40	X	X	.	X	X	X	X	X	X	X	X	X	X	X	0.01%
41	X	X	.	X	X	X	X	X	.	X	X	.	X	X	0.01%
42	X	X	.	X	X	.	X	X	X	X	.	X	X	X	0.03%
43	X	X	.	X	X	.	X	X	X	X	.	.	X	X	0.01%
44	X	X	.	X	X	.	X	X	.	X	.	X	X	X	0.01%
45	X	X	.	X	X	.	X	.	X	X	.	.	X	X	0.01%
46	X	.	X	X	X	X	X	X	X	X	X	X	X	X	0.01%
47	X	.	X	X	X	X	X	X	X	X	X	X	.	X	0.03%
48	X	.	X	X	X	X	X	X	X	X	X	.	X	X	0.01%
49	X	.	X	X	X	X	X	X	X	X	X	.	.	X	0.01%
50	X	.	X	X	X	X	X	X	.	X	X	.	X	X	0.01%
51	X	.	X	X	X	.	X	X	.	X	.	X	.	X	0.01%
52	X	.	X	.	X	X	X	X	.	X	.	.	X	X	0.01%
53	X	.	.	X	X	.	X	X	X	X	.	X	X	X	0.01%
54	X	.	.	X	X	.	X	X	X	X	.	.	X	X	0.03%
55	X	.	.	X	X	.	X	X	.	X	.	X	X	X	0.01%
56	.	X	X	X	X	X	X	X	X	X	X	X	X	X	0.01%
57	.	X	X	X	X	X	X	X	.	X	X	X	X	X	0.09%
58	.	X	X	X	X	X	X	X	.	X	X	X	.	X	0.01%
59	.	X	X	X	X	X	X	X	.	X	X	.	X	X	0.14%
60	.	X	X	X	X	X	X	X	.	X	X	.	.	X	0.04%
61	.	X	X	X	X	X	X	X	.	X	.	X	X	X	0.01%
62	.	X	X	X	X	.	X	X	.	X	.	X	X	X	0.03%
63	.	X	X	X	X	.	X	X	.	X	.	.	X	X	0.05%
64	.	X	.	X	X	X	X	X	.	X	X	X	X	X	0.01%
65	.	X	.	X	X	X	X	X	.	X	X	.	X	X	0.04%
66	.	X	.	X	X	X	X	X	.	X	.	.	X	X	0.01%
67	.	X	.	X	X	.	X	X	.	X	X	X	X	X	0.01%
68	.	X	.	X	X	.	X	X	.	X	X	.	X	X	0.08%
69	.	X	.	X	X	.	X	X	.	X	.	X	X	X	0.01%

70	.	X	.	X	X	.	X	X	.	X	.	.	X	X	0.03%
71	.	X	.	X	X	.	X	.	X	X	.	X	X	X	0.35%
72	.	X	.	X	X	.	X	.	X	X	.	X	.	X	0.14%
73	.	X	.	X	X	.	X	.	X	X	.	.	X	X	0.12%
74	.	X	.	X	X	.	X	.	X	X	.	.	.	X	0.01%
75	.	X	.	X	X	.	X	.	.	X	.	X	X	X	0.12%
76	.	X	.	X	X	.	X	.	.	X	.	X	.	X	0.07%
77	.	X	.	X	X	.	X	.	.	X	.	.	X	X	0.07%
78	.	.	X	X	X	X	X	X	.	X	X	.	X	X	0.01%
79	.	.	.	X	X	X	X	X	.	X	X	X	X	X	0.07%
80	.	.	.	X	X	X	X	X	.	X	X	.	X	X	0.34%
81	.	.	.	X	X	X	X	X	.	X	X	.	.	X	0.03%
82	.	.	.	X	X	X	X	X	.	X	.	.	X	X	0.01%
83	.	.	.	X	X	X	X	.	.	X	X	.	X	X	0.01%
84	.	.	.	X	X	.	X	X	.	X	X	X	X	X	0.05%
85	.	.	.	X	X	.	X	X	.	X	X	.	X	X	0.51%
86	.	.	.	X	X	.	X	X	.	X	X	.	.	X	0.03%
87	.	.	.	X	X	.	X	X	.	X	.	X	X	X	0.03%
88	.	.	.	X	X	.	X	X	.	X	.	.	X	X	0.09%
89	.	.	.	X	X	.	X	.	.	X	X	.	X	X	0.03%

Appendix W. Correlation between Schoenfeld residuals and follow-up time functions (Cohort 1)

		follow-up time	**log time***	**time squared#**	**Schoenfeld Residuals of alcohol use**
follow-up time	**r**	1.0000	0.9591	0.9813	-0.0039
	p-value	-	<0.0001¶	<0.0001¶	0.8742
log time*	**r**	0.9591	1.0000	0.8906	-0.0035
	p-value	<0.0001¶	-	<0.0001	0.8894
time squared#	**r**	0.9813	0.8906	1.0000	-0.0079
	p-value	<0.0001¶	<0.0001¶	-	0.7497
Schoenfeld Residuals of alcohol use	**r**	-0.0039	-0.0035	-0.0079	1.0000
	p-value	0.8742	0.8894	0.7497	-

r = Pearson correlation coefficient

¶ Indicates a P-value<0.05, *Log function of follow-up time, # Square function of follow-up time

Appendix X. Correlation between Schoenfeld residuals and follow-up time functions (Cohort 2)

		follow-up time	log time*	time squared#	Schoenfeld Residuals of FV intake
follow-up time	**r**	1.0000	0.9799	0.9892	-0.0413
	p-value	-	<0.0001¶	<0.0001¶	0.3369
log time*	**r**	0.9799	1.0000	0.9414	-0.0405
	p-value	<0.0001	-	<0.0001¶	0.3463
time squared#	**r**	0.9892	0.9414	1.0000	-0.0417
	p-value	<0.0001¶	<0.0001¶	-	0.3322
Schoenfeld Residuals of FV intake	**r**	-0.0413	-0.0405	-0.0417	1.0000
	p-value	0.3369	0.3463	0.3322	-

r=Pearson correlation coefficient

¶ Indicates a P-value<0.05, *Log function of follow-up time, # Square function of follow-up time

Curriculum Vitae

Name: Yuguang Kang

Post-secondary Education and Degrees:

Western University
London, Ontario, Canada
2018-2021 M.Sc.

Jilin University
Changchun, Jilin, China
2004-2007 M.Sc.

Changzhi Medical College
Changzhi, Shanxi, China
1998-2003 B.Sc.

Related Work Experience:

Research Analyst
London Health Sciences Centre/ICES
London, Ontario, Canada
2020-present

Graduate Research Assistant
Western University
London, Ontario, Canada
2018-present

Cardiologist
Shanghai No.3 People's Hospital in Shanghai Jiao Tong University School of Medicine

Shanghai, China

2007-2014

Honors and Awards: Western Graduate Research Scholarship 2018-2020

Funded Grants:

Project Title: Endothelial dysfunction in cardiovascular disease

- Co-investigator
- Primary Investigator: Professor Ping Yang, Jilin University, Jilin Province, China. Funded by: Ministry of Science and Technology of Jilin Province (200505173) Funds: 200,000 RMB

Presentations:

Poster Presentation: Optimal Antithrombotic Strategy for Patients with Atrial Fibrillation and Acute Coronary Syndrome: A Systematic Review and Meta-Analysis. Kang Y., Karp I., Anderson K. London Health Research Day; London, ON. March 30, 2019.

Publications:

Changes of Vascular Endothelial Cell Function in Patients with Viral Myocarditis by High Resolution Ultrasonography. Kang Y., Yang P., Yu J. Chinese Circulation Journal. 2008 Nov., 23(4) :764-8

Book Chapters:

Treatment of Complications of Cardiovascular Disease. Yang P. Liu B. Kang Y. 2008. Jilin Province Science and Technology Press. ISBN: 978-7-5384-1534-8