SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

School of Medicine and Surgery

PhD program in Public Health
XXXI Cycle

Curriculum in Epidemiology and Research on Prevention,
Diagnosis and Treatment Systems

DIETARY NON ENZYMATIC ANTIOXIDANT CAPACITY AND THE RISK OF CARDIOVASCULAR DISEASES – AN EPIDEMIOLOGICAL APPROACH

Surname HANTIKAINEN Name ESSI MARJATTA

Registration number 81106

Tutor: Prof. RINO BELLOCCO

Co-tutor: YLVA TROLLE LAGERROS, MD, PhD

Coordinator: Prof. GUIDO GRASSI

ACADEMIC YEAR 2018/2019

To my family

ABSTRACT

Cardiovascular diseases are the leading cause of premature death and disability in the world. A diet containing high amounts of plant-based foods has been associated with a reduced risk of cardiovascular diseases and the beneficial effect has been attributed to the antioxidants found in the foods. However, findings from randomized controlled trials on the role of antioxidant supplementation have been disappointing, reporting null results or even harmful effects. It has been suggested that antioxidants interact with each other to promote cardiovascular health. Therefore, the Non Enzymatic Antioxidant Capacity (NEAC) assay has been proposed, which measures the antioxidant potential of different dietary sources considering interactions between them.

This thesis aimed to further clarify the effect of dietary antioxidants on the risk of cardiovascular diseases, with particular interest in measuring NEAC from diet. The specific aims were to prospectively study whether dietary NEAC is associated with a lower risk of myocardial infarction, stroke and heart failure in subjects free from CVD or cancer. Four studies were conducted using data from two large Swedish cohorts. Multivariable Cox proportional hazard regression models were fitted to estimate hazard ratios (HRs) with 95% confidence intervals (CIs).

In the Swedish Women's Lifestyle and Health Cohort (n = 45,882), a higher baseline dietary NEAC was inversely associated with the risk of myocardial infarction (quintile 5 vs. quintile 1: HR: 0.60, 95% CI: 0.45-0.81, p for trend < 0.05) and heart failure (tertile 3 vs. tertile 1: HR: 0.63; 95% CI: 0.43-0.93; p for trend < 0.05) in young to middle aged women, whereas no association was found between dietary NEAC and stroke. In the Swedish National March Cohort (n = 34,543), dietary NEAC was inversely associated with the risk of overall (quartile 4 vs. quartile 1: HR: 0.77, 95% CI: 0.61-0.96; p for trend < 0.05) and non-fatal myocardial infarction (quartile 4 vs. quartile 1: HR: 0.72; 95% CI: 0.56-0.92; p for trend < 0.05), but not with fatal myocardial infarction. The association seemed to further be stronger in women compared to men.

To conclude, these findings support the hypothesis that a diet with high NEAC might protect from the development of myocardial infarction and heart failure and that the beneficial effect might be exerted through interactions between antioxidants. Whether this is true for stroke needs to be further investigated. Nevertheless, it is suggested to implement high amounts of antioxidant rich foods and beverages, such as fruits, vegetables, whole grains and tea, in the daily diet to lower the burden of cardiovascular diseases.

LIST OF SCIENTIFIC PAPERS

I. **Essi Hantikainen**, Marie Löf, Alessandra Grotta, Ylva Trolle Lagerros, Mauro Serafini, Rino Bellocco and Elisabete Weiderpass. Dietary non enzymatic antioxidant capacity and the risk of myocardial infarction in the Swedish women's lifestyle and health cohort. *European Journal of Epidemiology. 2018;33(2):213-221.*

II. **Essi Hantikainen**, Alessandra Grotta, Mauro Serafini, Ylva Trolle Lagerros, Olof Nyren, Weimin Ye, Luca Colarusso and Rino Bellocco. Dietary non-enzymatic antioxidant capacity and the risk of myocardial infarction: the Swedish National March Cohort. *International Journal of Epidemiology. 2018;47(6):1947-1955*.

III. **Essi Hantikainen**, Marie Löf, Alessandra Grotta, Ylva Trolle Lagerros, Mauro Serafini, Rino Bellocco and Elisabete Weiderpass. Dietary non enzymatic antioxidant capacity and the risk of stroke - the Swedish women's lifestyle and health cohort. *Submitted*

IV. **Essi Hantikainen**, Marie Löf, Alessandra Grotta, Ylva Trolle Lagerros, Mauro Serafini, Rino Bellocco and Elisabete Weiderpass. Dietary non enzymatic antioxidant capacity and the risk of heart failure - the Swedish women's lifestyle and health cohort. *Submitted*

RELATED PUBLICATIONS

I. Ylva Trolle Lagerros, **Essi Hantikainen**, Daniela Mariosa D, Weimin Ye, Hans-Olov Adami, Alessandra Grotta, Francesca Ghilotti and Rino Bellocco. Cohort profile: the Swedish National March Cohort. *International Journal of Epidemiology*. 2016;46(3):795-795e.

II. Ylva Trolle Lagerros, **Essi Hantikainen**, Karl Michaëlsson, Weimin Ye, Hans-Olov Adami and Rino Bellocco. Physical activity and the risk of hip fracture in the elderly: a prospective cohort study. *European Journal of Epidemioly. 2017;32(11):983-991.*

III. Lois Veen, **Essi Hantikainen**, Rino Bellocco, Mauro Serafini, Weimin Ye, Alessandra Grotta, Ylva Trolle Lagerros. Dietary Non Enzymatic Antioxidant Capacity and the risk of Osteoarthritis in the Swedish National March Cohort. *Submitted*

CONTENTS

LIST OF ABBREVIATIONS

AP	Proportion attributable to interaction
BMI	Body mass index
CVD	Cardiovascular disease
DASH	Dietary Approaches to Stop Hypertension
FFQ	Food frequency questionnaire
FRAP	Ferric reducing antioxidant power
GBD	Global burden of disease
HAT	Hydrogen Atom Transfer
HF	Heart failure
HR	Hazard ratio
ICD	International coding of disease
MET	metabolic equivalent of task
MI	Myocardial infarction
NEAC	Non-enzymatic antioxidant capacity
ORAC	Oxygen radical absorbance capacity
PAR	Population attributable risk
PH	Proportional hazard
PIN	Personal identity number
RERI	Relative excess risk due to interaction
TEAC	Trolox Equivalent Antioxidant Capacity
TRAP	Total radical trapping parameter
S	Synergy index
SET	Single Electron Transfe
SNMC	Swedish National March Cohort
WLHC	Swedish Women's Lifestyle and Health Cohort

1 INTRODUCTION

Although the global burden of cardiovascular disease (CVD) has decreased in the past decade, CVD remains the leading cause of death and disability in the world [1]. In Europe one of every four deaths and in the United States one of every three deaths occurs due to CVD [2, 3]. As diet explains about one third of global mortality, diet is a key modifiable risk factor and target in the development and prevention of CVD [4]. In fact, diet has been widely investigated and different approaches, such as studying single nutrients, foods and dietary patterns, have been used to identify a healthy diet. Nevertheless many questions remain unanswered.

There is great evidence that a diet containing high amounts of plant-based foods, such as fruits, vegetables and whole grains, reduces the risk of CVD. This protective effect has been attributed to dietary antioxidants found in foods, which are able to scavenge free radicals and to reduce oxidative stress [5]. However, findings from observational studies investigating single antioxidants and the risk of CVD have cast doubt on the association between antioxidants and CVD [6]. What has not been considered in these studies is, that antioxidants interact with each other to promote cardiovascular health. Therefore, a single antioxidant approach might not be suitable to study the effect of dietary antioxidants on the risk of CVD.

This thesis aimed to further clarify the effect of dietary antioxidants on the risk of selected cardiovascular outcomes, i.e. myocardial infarction, stroke and heart failure, with particular interest to investigate the effect of dietary non-enzymatic antioxidant capacity (NEAC), which measures the antioxidant potential of different dietary sources considering interactions between them. To accomplish the aim four observational studies were conducted using data from two large Swedish cohorts.

2 BACKGROUND

2.1 Cardiovascular diseases

Cardiovascular diseases (CVDs) are a group of diseases affecting the heart and blood vessels, with main manifestations being ischemic heart disease (ISH), cerebrovascular disease and peripheral vascular disease [7]. CVDs are the number one cause of death worldwide, with IHD and stroke remaining the leading causes of premature death and disability, both globally and in each world region. According to the Global Burden of Disease Study 2015 (GBD) there were an estimated 422.7 million prevalent CVD cases and 17.9 million CVD deaths, which accounted for one third of all deaths in the world [1].

Age-standardized prevalence of CVDs vary greatly among countries, with lowest rates in Singapore, Japan, New Zealand, followed by Western Europe and the United States, and highest rates in West and East Africa, Iran and Oman. Although prevalence of CVDs is declining in many high and some middle-income countries, mortality rates seem to have plateaued, especially in Western Europe and the United States [1]. Based on the report conducted by Townsend and colleagues in 2016 [2], still 3.9 million deaths are caused by CVDs every year in Europe, with more than 1.4 million people dying before the age of 75 years . In the United States an estimated number of 840,000 annual CVD deaths occurred in 2016 [8].

Because of the ageing of the population, the epidemiological transition of low- and middle-income countries from infectious to chronic diseases, as well as the increase in modifiable risk factors, the burden of CVD is expected to increase. This trend is worrisome and demonstrates the importance of investment in prevention and treatment of CVD worldwide [1].

2.1.1 Myocardial infarction

Ischemic heart disease (IHD) is the primary cause of mortality and morbidity in the world [1], with myocardial infarction (MI) being the first manifestation of IHD. The pathological characteristic of MI is cell death due to prolonged ischemia, which causes an imbalance between oxygen supply and demand in the myocardium [9]. Main cause for MI is the rupture of an atherosclerotic lesion in a coronary artery, which leads to the formation of a thrombus that plugs the artery leading to ischemia by stopping blood flow to the heart [7].

In 2015 there were around 110 million prevalent cases and 8.9 million fatal cases of IHD worldwide [1]. After the age of 40 years death rates increase steeply, from an estimated 33 deaths per 100,000 for those aged 40-44 years, up to 2,671 per 100,000 for those above the age

of 80 years. When looking at MI separately, around 7.3 million acute cases were recorded globally in 2015 [1]. Between 1990 and 2010 MI incidence rates have decreased in all age groups, as well as in men (from 222.7 to 195.3 per 100,000) and women (136.3 to115.0 per 100,000) [10]. However, within a year of first MI more women than men will die (26% of women, 19% of men), and within a five year period an estimated 47% of women and 36% of men will die, have heart failure or stroke, regardless of age [11].

2.1.2 Stroke

Stroke is the second leading cause of death in the world and responsible for long-term disability and dependency on patient care [12, 13]. Stroke is characterized by an acute focal injury of the central nervous system caused by inadequate blood flow [14]. Depending on the site of the injury, stroke can be categorized into different subtypes. Cerebral infarction, also known as ischemic stroke, occurs in 85% of stroke cases and is caused by local injury leading to the formation of thrombi and subsequent occlusion of the cerebral vessels or due to occlusion of the cerebral vessels by a circulating blood clot [12, 15]. A local breaking of the cerebral blood vessels causing hematoma will lead to hemorrhagic stroke, the most devastating type of stroke with high mortality rates and leaving 75% of the patients disabled and diseased [16]. Subarachnoid hemorrhage is caused by a bleeding into the subarachnoid space surrounding the brain and spinal cord [17].

Although age-standardized mortality rates and prevalence of stroke have decreased, the overall burden of stroke remains high [13]. In 2016 the Global Burden of Disease Report estimated 80.1 million prevalent cases of stroke in the world, of which 41.1 million occurred in women and 39.0 million in men. In addition, 5.5 million deaths due to stroke were reported, of which 2.8 million were due to hemorrhagic and 2.7 million due to ischemic stroke. Age-specific stroke incidence rates were comparable between men and women under the age of 55, but significantly higher for men compared to women between the ages of 55-75 years [18].

2.1.3 Heart failure

Heart failure (HF) is a major public health issue affecting around 26 million people worldwide [19]. It is a complex and progressive disorder caused by the incapability of the heart to generate sufficient cardiac output to meet the bodies demand. In 70% of the cases HF is caused by impaired ventricular contraction [7], which generally is a consequence of IHD and MI. After MI the myocardium tries to adapt to the increased wall stress, which involves changes in the

structure and function of cardiac myocites, a process also known as cardiac remodeling leading to HF [20].

Recent evidence is indicating that the incidence of HF seems to have stabilized, or even decreased in women [21]. Nevertheless, prevalence of HF is rising due to an ageing population and improved treatments of CVDs. In the US and Europe more than 80% of death and prevalent cases of HF can be attributed to individuals older than 65 years. Moreover, HF has a great impact on quality of life and is often worse than for other chronic diseases [19]. Especially women show poorer social and physical functioning compared to men after HF diagnosis [22].

2.1.4 Risk factors

MI, stroke and HF share common risk factors, which can be divided into non-modifiable and modifiable ones. Next to some non-modifiable risk factors, which are age, sex and family history of CVD [23], other well established modifiable risk factors are preclinical conditions and lifestyle behaviors, such as diabetes, smoking, obesity, physical activity, diet and alcohol consumption [24].

Several studies have tried to identify the population attributable risk (PAR), also known as the population attributable fraction [25], for CVDs with focus on modifiable risk factors. The PAR is defined as the proportion of the incidence of a disease in the population that is due to exposure. In other words, it is the incidence of a disease in the population that would be eliminated if the exposure were removed [26]. The PAR is calculated with the following formula [27]:

$$PAR = \frac{P_e(RR - 1)}{1 + P_e(RR - 1)} \tag{2.1}$$

where,

P_e = Probability of the exposure to the risk factor

RR = Relative risk of the disease in exposed versus unexposed individuals

The PAR is usually expressed as a percentage, which can be calculated by dividing the PAR by the incidence in the total population and by multiplying the product by 100 [26].

Findings on myocardial infarction

The INTERHEART study, a large case-control study including subjects from 52 countries, estimated PAR% for several risk factors and found lipid disturbance (49.2%) and smoking (35.7%) to be the strongest predictors for MI, followed by psychosocial factors (32.5%), abdominal obesity (20.1%), hypertension (17.9%) and diabetes (9.9%). Daily consumption of fruits and vegetables (13.7%), moderate to strenuous physical activity (12.2%), as well as moderate alcohol consumption (6.7%), were shown to protect from MI. These findings were consistent among men and women and across different age groups, ethnicities and geographic regions [27].

Findings on stroke

The INTERSTROKE study, a large case-control study including subjects from 32 countries, reported hypertension to be the leading risk factor for stroke with a PAR% of 47.9%, followed by lipid disturbance (26.8%), waist-to-hip ratio (18.6%), psychosocial factors (17.4%), smoking (12.4%), cardiac cause (9.1%), high alcohol consumption (5.8%) and diabetes (3.9%). Being physically active (35.8%) and following a healthy diet (measured by adherence to the Alternative Healthy Eating Index) (23.2%) were protective [28].

Findings on heart failure

In the NHANES-I study, a large-prospective US cohort study, risk factors and their attributable risk estimated in PAR% for heart failure were reported. They found that 61.6% of HF is attributable to IHD, followed by smoking (17.1%), hypertension (10.1%), physical inactivity (9.2%), education (8.9%), obesity (8.0%), diabetes (3.1%) and valvular heart disease (2.2%). Diet was not considered in this study [29].

Together all modifiable risk factors are estimated to account for more than 90% of the total risk of developing MI and stroke and around 40% of the total risk of HF. Therefore it is crucial to identify sustainable lifestyle interventions to reduce the burden of CVD worldwide [24].

2.2 Oxidative stress

2.2.1 Oxidative stress and the antioxidant defense system

Oxidative stress is a condition characterized by an imbalance between free radicals and the bodies own antioxidant defense system, which leads to cell and tissue damage [30]. Free radicals, such as reactive oxygen species (ROS), are highly reactive atoms or molecules, which try to capture electrons from other molecules to neutralize themselves [30, 31]. This mechanism causes a chain reaction that leads to the formation of further free radicals, until the subsequent free radicals are deactivated [31]. The formation of free radicals occurs in all aerobic cells and plays an important role in ageing and age related diseases [30]. To protect the cells from damage through free radicals, a complex antioxidant defense system has evolved that consists of endogenous and exogenous components, which interact with each other to neutralize free radicals [31].

Endogenous antioxidants, which are the bodies own antioxidants, include antioxidant enzymes, metal binding proteins and copper ions, and are critical for maintaining cellular health [31]. However, if the human body is exposed to external stressors, such as infections, toxins, radiation and smoking, the endogenous antioxidant defense system is not enough and needs support from exogenous antioxidants found in the diet [30, 31].

2.2.2 Oxidative stress and cardiovascular disease

Oxidative stress has been associated with development of atherosclerosis, which is the main precursor for CVDs [32]. It has been suggested that free radicals are involved in the oxidation of lipids, membranes, proteins and DNA [5]. In association with atherosclerosis, the combination of free radicals and elevated and modified circulating low density lipoproteins induces oxidation of circulating lipoproteins and their uptake into the artery wall, which contributes to endothelial dysfunction [33]. In the endothelium oxidised lipoproteins are converted to highly oxidised LDL. These are taken up by magrophages to form foam cells, which in turn leads to the formation of fatty streaks (**Figure 2.2.1**). A subsequent lesion of the vascular endothelial cell activates the release and formation of thrombi, yealding to occlusion of the vessel at the site of the lesion or elswhere in the coronary or cerebral artery, causing myocardial infarction or stroke, respectively [7, 34]. In heart failure oxidative stress is responsible for cardiac remodeling by inducing hypertrophic signaling, apoptosis and necrosis of the cardiac cells [35].

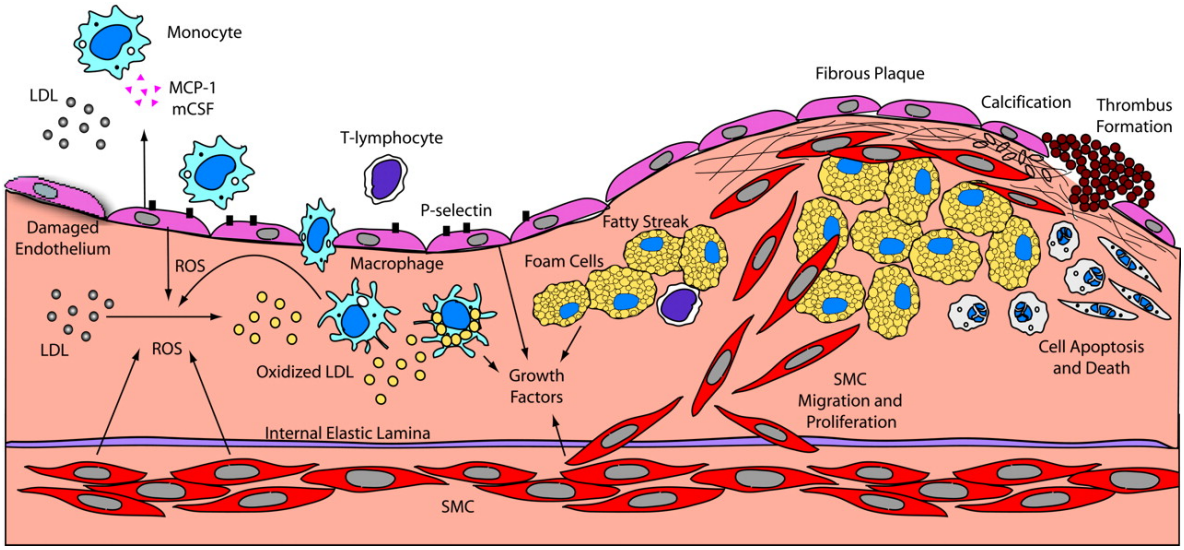

Figure 2.2.1. Development of atherosclerosis. ROS produced by endothelial cells, smooth muscle cells; macrophages oxidize LDL in the subendothelial space; at the sites of endothelial damage, initiating events that terminate in the formation of a fibrous plaque; rupture of fibrous plaque leads to thrombus formation and occlusion of the vessel. Reproduced with permission from Madamanchi et al. (2005).

2.3 Diet

Diet plays a key role in the prevention of CVDs and its relevance has been detected already in the early 1900, when high cholesterol intake was shown to increase the risk of atherosclerosis in rabbits. Despite huge efforts in nutritional research in the past decades, our knowledge of cardiovascular effects of diet is still very limited [4]. There is, however, growing evidence that the consumption of certain foods lead to a reduction of oxidative stress and atherosclerosis and this protective effect has been attributed to the antioxidants found in the foods [31, 36].

This section covers current evidence from observational studies and randomized controlled trials investigating the role of diet, dietary patterns and antioxidants in relation to CVD risk.

2.3.1 Diet, dietary patterns and cardiovascular disease

Findings from several prospective studies have linked an antioxidant rich diet to lower risk of MI, stroke and HF, showing great evidence for a protective effect from a higher intake of plant-based foods [37-39]. A recent meta-analysis concluded, that with each 200g/day increment of fruits and vegetables the risk of IHD, stroke and overall CVD was reduced by 8-16%, 13-18% and 8-13%, respectively [40]. Further, a systematic-review concluded that a higher intake of

whole grains, cereal fiber and bran might lower risk of CVD [41]. In addition, consumption of coffee, tea and chocolate might protect from the development of CVD [42, 43] .

Besides some specific foods there is evidence from both observational studies and randomized controlled trials linking several dietary patterns to a lower risk of CVD. For example, the Dietary Approaches to Stop Hypertension (DASH) diet, which is rich in fruits, vegetables and low-fat dairy products and low in total and saturated fat intake [4], was shown to protect from MI, stroke and HF [44]. Similarly, the Mediterranean diet, which is characterized by high intake of fruits, vegetables, whole grains, plant-based fats (olive oil, nuts) and low intake of meat [4], has been associated with a lower risk of CVD incidence and mortality [37, 45]. Both diets might reduce CVD risk by down regulating low-grade inflammation and by improving other lifestyle risk factors [46].

2.3.2 Antioxidants and cardiovascular disease

As mentioned earlier, dietary antioxidants are considered exogenous antioxidants, which support the bodies own antioxidant defense system to scavenge free radicals and to reduce oxidative stress. Especially certain antioxidants, i.e. vitamin C, vitamin E, beta-carotene, polyphenols, and their relation to CVDs have been widely investigated [31, 47]. In addition, vitamin supplementation as a potential therapeutic intervention in the prevention of CVDs has reached considerable interest [48].

Evidence from observational studies

Previous observational studies investigating the effect of antioxidants coming from both natural sources and supplements reported inverse associations with the risk of CVD in some, but not all studies [49-53]. Vitamin C is a water-soluble vitamin that naturally occurs in citrus fruits, tomatoes and potatoes. A meta-analysis comprising 15 cohort studies [54] reported a significant inverse association with vitamin C intake and IHD, for both total and dietary vitamin C but not supplemental vitamin C. Vitamin E is a fat-soluble vitamin consisting of a mixture of tocopherols and tocotrienols and is mainly found in vegetal oils. For vitamin E an inverse association was found for total, dietary and supplemental intake and IHD risk. Beta-carotene, a fat-soluble vitamin and mainly found in carrots, tomatoes and dark green leafy vegetables [31], has not been associated with the risk of IHD [54]. Polyphenols are the most abundant antioxidants found in the diet and mainly occur in fruits, vegetables, plant-derived beverages

(tea, coffee, red wine), cereals and legumes [55]. A meta-analysis including 14 cohort studies found flavonoid intake to be inversely associated with CVD risk [56].

Evidence from randomized controlled trials

Findings from clinical trials on vitamin supplementation in the prevention of cardiovascular diseases are inconclusive, reporting mainly null results or even harmful effects. In the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study male smokers aged 50 to 69 years were assigned to either receive vitamin E (50mg), beta-carotene (20mg), both agents or placebo for a period of five to eight years. A marginal effect was found between vitamin E and incidence of fatal IHD in subjects without prior myocardial infarction, but no effect was seen for nonfatal MI. In addition, no association was found between beta-carotene and major cardiac event [57]. In the Women's Health Study consumption of natural vitamin E (600IU) every other day over a mean follow-up time of 10.1 years was not associated with the risk of MI and stroke, however, an inverse association was found for cardiovascular death (RR: 0.76; 95% CI: 0.59-0.98; p for trend = 0.03) [58]. In the Physicians Health Study II long-term supplementation with synthetic vitamin E (400IU/every other day), vitamin C (50mg/day) and multivitamins was not associated with the risk of developing major cardiovascular events, such as MI and stroke, in young to middle aged men [59, 60]. Further, some studies have raised concern about an increased risk of HF related to vitamin E supplementation in subjects who are at high cardiovascular risk. In the HOPE and HOPE-TOO trial long-term intake of natural source vitamin E (RRR-tocopheryl acetate) (400IU/day) over a mean follow-up of seven years was investigated in patients with vascular disease or diabetes mellitus. No association was found with major cardiovascular events, but higher rates were seen for heart failure and hospitalizations for heart failure [61]. Similarly, in the GISSI-Prevention study, an open-labeled randomized controlled trial with a follow-up of 3.5 years, vitamin E intake (300mg/day) was associated with an increased risk of HF in subjects with prior MI [62]. A first meta-analysis published in 2003 analyzed seven randomized trials of vitamin E treatment and, separately, eight of beta-carotene treatment, including 1000 individuals. No association was found between vitamin E intake and the risk of CVD but a slightly increased risk with beta-carotene [63]. Two more recent meta-analyses have been conducted since. One study comprising 15 clinical trials including 188,209 individuals found no association between synthetic or natural vitamin supplementation and CVD risk. However, the study was not able to account for potential

differences of supplement dose and follow-up times [6]. The other study comprising 50 clinical trials including 294,478 individuals was able to account for subgroup effects by type of prevention, type of vitamins and antioxidants, type of cardiovascular outcomes, duration of treatment and funding source, amongst others [64], and concluded that vitamin supplementation has no effect on the incidence of major cardiac events, MI, stroke and cardiac death. Based on these findings vitamin supplementation is not recommended for the prevention of CVDs [6, 64].

2.4 Non Enzymatic Antioxidant Capacity

There is growing evidence that single nutrients have limited effects on chronic diseases compared to whole foods [4]. In fact, antioxidants interact cumulatively and synergistically with each other to exert their full antioxidant capacity in the human body [36, 65]. Therefore, a measure called non-enzymatic antioxidant capacity (NEAC) has been proposed, which captures the antioxidant potential of the whole diet taking interactions between them into account.

In 2004 the First International Conference on Antioxidant Methods was held in Orlando, Florida, to address analytical issues to analyze Antioxidant Capacity of foods. The antioxidant system is very complex, which brings several challenges with it to find an appropriate assay to measure antioxidant capacity. First, there are at least four general sources of endogenous and exogenous antioxidants, which can either be enzymes (e.g. superoxide dismutase, glutathione peroxidase), large molecules (e.g. albumin, ferritin), small molecules (e.g. uric acid, ascorbic acid, tocopherols, carotenoids, phenols) and some hormones (e.g. estrogen, melatonin). In addition, both antioxidants and oxidants have different chemical and physical characteristics, sometimes showing multiple mechanisms depending on the reaction system. Therefore, no single assay exists, that is able to accurately reflect all antioxidants and sources of free radicals in a complex system [66].

Non-enzymatic antioxidant capacity, also known as Total Antioxidant Capacity, measures the capability of a given test solution to scavenge free radicals. NEAC is a complex measure integrating many variables, such as redox potential of the nutrient compounds, cumulative and synergistic interactions between them, as well as the kind of stress found in the test environment [65]. It has further been stated, that NEAC should be seen as a concept that reflects the multiple aspects of redox interactions, rather than an analytical technique. It should

help us to better understand how dietary factors are involved in complex pathological processes [65].

2.4.1 Chemical assays to measure NEAC

There are two major mechanisms how antioxidants can deactivate free radicals, i.e. the Hydrogen Atom Transfer (HAT) and the Single Electron Transfer (SET). Based on these mechanisms different assays have been developed to measure antioxidant capacity [66]. The HAT-based methods measure the ability of antioxidant to neutralize free radicals by hydrogen donation. The SET-based methods measure the ability of an antioxidant to transfer an electron to another compound [66]. The assays differ in characteristics, such as molecular target, endpoint and physiological relevance. The most used assays and databases are briefly described below.

Ferric Reducing Antioxidant Power (FRAP)

FRAP is a method based on the SET reaction mechanism. It is a simple and highly reproducible assay using FE(II) as a standard, that measures the ability of plasma samples to reduce ferric-tripyridyltriazine complex (TPTZ-Fe3+) to TPTZ-Fe2+, the "reducing power". However, it is not able to measure lipophilic antioxidants and thiol groups [65].

Total Radical Trapping Parameter (TRAP)

TRAP is a method that uses the HAT reaction mechanism and measures the ability of plasma to scavenge peroxyl radicals, the "chain braking" antioxidant activity of plasma. TRAP is specifically suitable to assess total antioxidant capacity of plasma samples but it should not be used to assess NEAC of lipophilic antioxidants [65].

Oxygen radical absorbance capacity (ORAC)

ORAC constitutes of the same chemical principle as TRAP using the HAT reaction mechanisms [66], but uses the area under the curve of fluorescence over time to measure plasma NEAC. The main limitation by this method is the interference with proteins, which leads to an underestimation of the contribution of the main chain braking antioxidant. In addition, lipophilic antioxidants cannot be directly measured with this assay [65].

NEAC databases

In 2003 and 2006 Pellegrini and colleagues published the first antioxidant food database, providing NEAC values based on three different antioxidant assays (FRAP, TRAP and Trolox Equivalent Antioxidant Capacity (TEAC)) for 104 commonly consumed food items in Italy [67, 68]. In 2007 the U.S. Department of Food and Agriculture followed with an antioxidant database and presented the ORAC of Selected foods report for 277 food samples [69]. In 2010, Carlson and colleagues developed a comprehensive database providing NEAC values using the FRAP assay for more than 3,100 food items consumed worldwide [70]. As all databases have used different methods to assess NEAC, the comparison of NEAC values for food items or groups is difficult. However, a food that has a high antioxidant value in one assay will likely have a high antioxidant value using another assay. Therefore, although the exact values will be different, the ranking of the foods will be similar no matter which assay is used [70].

2.4.2 NEAC and cardiovascular disease – current evidence

NEAC has recently gained great attention and has been used to assess association between dietary and plasma NEAC and CVD risk factors. For example, a diet based on high NEAC has been associated with improved endothelial function, a less atherogenic blood profile and a reduction of inflammatory biomarkers associated with cardiovascular diseases [71-73]. A recent review of 16 studies reported higher dietary NEAC to be further inversely associated with other CVD risk factors, such as C-Reactive protein, blood pressure, and waist circumference [74].

To date, few observational studies examined the association between dietary NEAC and CVD risk. Two studies found a diet with high NEAC to be associated with lower MI risk in men and women [75, 76]. For stroke, the evidence is still less clear. Three studies found a protective effect, however, results differ between types of stroke and sex [77-79]. Moreover, one study found no association at all [80]. In addition, only one study has examined the effect of dietary NEAC on the risk of heart failure, reporting an inverse association in women [81]. Evidence for an association between dietary NEAC and the risk of CVD is therefore still limited. In addition, all studies have used different antioxidant food databases and chemical assays to assess dietary NEAC.

3 AIMS

The overall aim of this thesis was to examine the relationship between the overall antioxidant capacity of the diet and the risk of several cardiovascular outcomes in two large Swedish cohorts.

The specific aims of the thesis are:

1. To investigate whether a higher dietary non-enzymatic antioxidant capacity is associated with the risk of myocardial infarction in men and women (Paper I+II).
2. To examine the association between dietary non-enzymatic antioxidant capacity and the risk of stroke in women (Paper III).
3. To study the association between dietary non-enzymatic antioxidant capacity and the risk of heart failure in women (Paper IV).

4 MATERIAL

4.1 Study design and populations

The studies included in this thesis were conducted based on two large Swedish cohorts with a long follow-up. Both cohorts were initially initiated to study the association between lifestyle factors and chronic diseases by collecting self-reported information on lifestyle habits and disease history with extensive baseline questionnaires. This section will introduce the cohorts and present some important findings in relation to the topic of this thesis.

4.1.1 Swedish Women's Lifestyle and Health Cohort

The Swedish Women's Lifestyle and Health Cohort (WLHC) was set up between 1991 and 1992 with the initial aim to investigate the associations between lifestyle factors, cancer and cardiovascular diseases among young and middle aged women. To do so, women between the age of 30 and 49 years residing in the Uppsala Health care region were randomly selected from the Swedish Population Registry at Statistics Sweden and invited to participate in the Swedish component of the Women's Lifestyle and Health Cohort. An invitation letter including a baseline questionnaire was mailed to 96,000 women, of whom 49,259 (51%) agreed to participate in the cohort. In addition, a follow-up questionnaire was sent out in 2003 to all women, who were still alive and residing in Sweden, with a response rate of 73% [82]. However, the follow-up questionnaire did not include any food frequency questionnaire.

Several studies within the WLHC have been published on different exposures and lifestyle related diseases, such as oral contraceptive use, BMI, exposure to UV radiation, alcohol consumption and dietary intake, on the risk of different cancer sites and overall- and cause specific mortality. CVDs have been investigated to a lesser degree. For example, fruit and vegetable intake as well as healthy dietary patterns (Healthy Nordic Dietary Index) have been studied, however no association was found with the risk of total cancer or CVDs [83, 84]. In addition, coffee consumption was associated with a lower risk of all-cause mortality, but not with cancer or cardiovascular mortality [85].

4.1.2 Swedish National March Cohort

The Swedish National March Cohort (SNMC) was established in 1997 during a four-day national fundraising event, which was organized by the Swedish Cancer Society. The event took place in around 3600 cities and villages throughout the country. The cohort was initially designed to investigate associations between lifestyle factors and chronic diseases. During the event, participants were asked to fill out a 36-page questionnaire, which covered detailed

questions about lifestyle factors and disease history. Due to the nature of the event, the total number of participants could not be assessed. In total, 43,880 participants returned a completed questionnaire [86].

The SNMC has been used to study different lifestyle factors and lifestyle related diseases, like physical activity, sleep, snus consumption in relation to different health outcomes, such as cancer, CVD and mortality. Diet has been studied to a lesser degree. For example, a higher dietary NEAC was associated with a lower risk of stroke and hip fracture [79, 87]. Other lifestyle factors, such as physical activity and sleep patterns, were associated with a lower risk of heart failure and specific types of CVD, respectively. The cohort has further been part of several consortia and has contributed to projects in the National Cancer Institute Cohort Consortium and the European Cohort Consortium [86].

4.2 Assessment of exposure variables

In nutritional epidemiology we are interested in the average intake of a given nutrient. The most widely used method to obtain self-reported information on dietary intake is the Food Frequency Questionnaire (FFQ). In all four studies the exposure of interest was dietary NEAC, which was assessed by combining information from baseline FFQs with two different antioxidant food databases. In this section the FFQs and the antioxidant food databases will be described in more detail.

4.2.1 Food Frequency Questionnaires

In the WLHC a previously validated [88], semi-quantitative FFQ was send out at baseline to all women participating in the study. The questionnaire covered information on frequencies and portion sizes for 88 foods and beverages consumed during the past year preceding enrollment and consisted of two sections. The first section covered foods consumed in Sweden on a regular base, such as milk, cheese, bread, coffee and alcohol. For these food items information on standard portion sizes were reported in the questionnaire and subjects were able to report consumption frequencies with open-ended response options. The second section of the FFQ covered 59 commonly eaten food items with eight response options: never/seldom; 1–3 times/month; 1, 2, 3–4 or 5–6 times/week; 1, 2 or 3 times/day. Subjects could further choose between the portion sizes small, medium and large. If subjects had left the information for the portion size blank, we assigned the smallest portion size to be able to keep the subjects in the study rather than excluding them. This assumption was justified based on a study showing that portion size has little effect on the relative contribution to the

total food or nutrient intake. This is a consequence of the smaller between-person variation arising from portion size compared to the between-person variation from frequency [89]. This FFQ was used in studies I, III and IV.

In the SNMC a similar, previously validated 85-item FFQ was collected at baseline [90], assessing dietary habits during the past year preceding enrolment. However, the portion sizes were fixed based on the standard portion size throughout the whole questionnaire. In addition, the FFQ used only preset response options, which were 0 to 7 times or more per day for commonly consumed foods and beverages, such as milk, cheese, bread, juice, coffee, or 0,1–3 times/month; 1–2, 3–4, 5–6 times/week; 1, 2, 3 times/day for other commonly eaten foods in Sweden. This FFQ was used in study II.

For all four studies, frequency variables were standardized into daily consumption. Missing values on frequencies were interpreted as null intakes [91], which has been shown to be a reasonable assumption used in nutritional epidemiology [92]. With both FFQs we used the Swedish National Food Administrative database to translate dietary information to energy and nutrient intake [93].

4.2.2 Antioxidant food databases

From the antioxidant food databases presented earlier, two were chosen for the assessment of dietary NEAC for our studies.

The antioxidant food database published by the Institute of Nutrition Research, University of Oslo, by Carlsen et al. [70] provides "*in vitro"* measurements of FRAP values (expressed per 100g of single foods) for more than 3100 foods and beverages. As stated previously, the original FRAP assay does not measure lipophilic antioxidants. However, for this database the FRAP assay was optimized to detect both, lipophilic and hydrophilic antioxidants [70].

This database was used for studies I, III and IV, which were conducted within the WLHC. All relevant items from the antioxidant food database were categorized according to the food groups or items used in the FFQ from the WLHC. We then computed median FRAP values for each category. If a category consisted of several food items (e.g. category "apple and pear" contains both, apple and pear; category "berry" consists of different berries), we used information on typical consumption percentages of different foods in Sweden to assign weights to each food item.

For study II, which was conducted within the SNMC, the antioxidant food database created by Pellegrini et al. was used [67, 68]. The database provides food item specific FRAP and TRAP values expressed per 1 kg fresh weight of single foods to compute daily dietary NEAC, based on mean values of 156 commonly consumed foods in Italy [67, 68]. The database already provides single values for each food item, which were computed based on the mean of three samples per item. Similarly as described for the antioxidant food database by Carlsen et al., we assigned weights to compute group specific FRAP and TRAP values if a food category in the FFQ consisted of more than one food item. The database of Pellegrini et al. was additionally used in studies I, III and IV as a secondary database to conduct sensitivity analyses by comparing different databases and different chemical assays to assess dietary NEAC.

To combine the information from the antioxidant food databases with the information reported in the FFQs used in the WLHC and the SNMC, all FRAP and TRAP values were multiplied by the consumption frequencies reported in the FFQs, by taking the portion size into account, and summed up to obtain subject specific total daily dietary NEAC at baseline.

4.2.3 Energy adjustment

When using dietary data in epidemiological studies it is crucial to account for total energy intake [94]. Most nutrients are associated with total energy intake, either because of their direct contribution to energy intake or because of higher consumption of energy, which in turn leads to a higher intake of all nutrients in general. Similarly, as nutrient intake and energy intake are computed from the same foods, measurement error from nutrient assessment is highly correlated with measurement error from energy assessment. Therefore, by controlling for variation in energy intake we are able to reduce measurement error of nutrient intake [94]. In addition, total energy intake is to a large extent a consequence of variations of physical activity, body size and metabolic efficiency. Therefore, they will likely cause between person variation, which is unrelated to the disease risk and thus extraneous. By adjusting for energy intake we can better control for extraneous variation induced by those factors [94].

Several models to adjust for energy intake can be used, such as the nutrient residual model, the nutrient density model, the energy partition model and the standard multivariable model, which are shortly presented in **Table 4.2.1** [92]. As we are often interested in categorizing nutrient variables, it should be noted, that the interchangeability of the models does not apply anymore, when categorical nutrient analyses are conducted. In fact, it has previously been

shown that in the categorical analysis of nutrient variables, statistical power is higher when the residual or the nutrient density model are used [92]. For studies I to IV the nutrient residual model was chosen to adjust nutrient variables for energy intake.

Table 4.2.1. Risk models for addressing the correlations of specific nutrients with total energy intake in epidemiological studies, adapted from Willet et al. (1997), Adjustment for total energy intake [94].

Model	*Relation of the model expressed*
Model 1 (Nutrient Residual model)	Disease risk = β_1 nutrient residual*
Model 2 (Nutrient Residual model + Energy intake)	Disease risk = β_1 nutrient residual* + β_2 total energy
Model 3 (Standard Multivariable model)	Disease risk = β_1 nutrient + α (total energy)
Model 4 (Energy partition model)	Disease risk = $(\alpha+\beta_1)$ nutrient + α (energy from non-nutrient sources)
Model 5 (Multivariable Nutrient Density model)	Disease risk = β_3 nutrient/total energy + β_4 total energy

*Nutrient residual refers to the residual from the regression of a specific nutrient on total energy intake

With the nutrient residual model the individual nutrient intake is regressed on the total energy intake [94]. The distributions of energy and nutrient intake are usually skewed towards extreme values. In addition the variation of nutrient intake, and therefore also the residuals, are larger at higher energy intake. It is thus recommended to do the energy adjustment after using log transformation of dietary variables, which creates residuals with a more constant variance across levels of total energy intake [95]. These residuals (a) represent the difference between each individual's actual intake and the predicted intake by their total energy intake, as presented in **Figure 4.2.1**. Because residuals have a mean of zero and include also negative values, it is desirable to add a constant to make the values more readable [94], which for studies I to IV was chosen to be the predicted nutrient intake for the mean energy intake of

the study population (b) (**Figure 4.2.1**) [94]. After, the anti-log can be taken to transform the energy-adjusted nutrient intake back into a familiar scale [95].

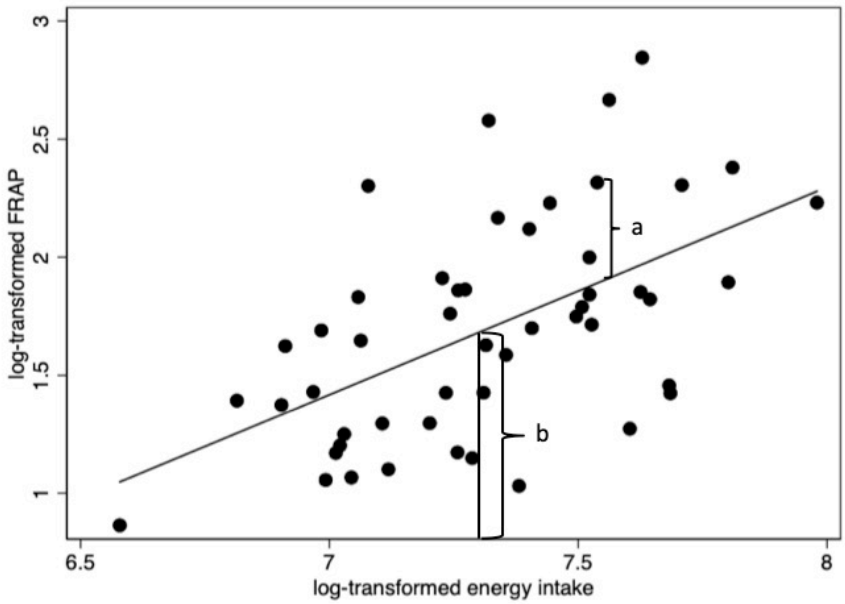

Figure 4.2.1. Energy-adjusted FRAP= a + b, where a=residual for subject from regression model with FRAP as the dependent variable and total energy intake as the independent variable and b= the expected FRAP for an individual with mean energy intake; FRAP and energy intake are presented on the log-transformed scale. Reproduced based on the example of Willet et al. (1997) for calorie adjusted nutrient intake; using the data from the Swedish Women's Lifestyle and Health Cohort with a random subsample of 50 women.

When applying the energy adjustment to dietary FRAP values obtained from the antioxidant food database from Carlsen et al. and using the data from the WLHC, the variation of dietary FRAP was reduced to some degree (**Figure 4.2.2**) (FRAP before adjustment: mean = 6.17, SD = 3.16; shaded area; FRAP after adjustment: mean = 6.00, SD = 2.83, non-shaded area).

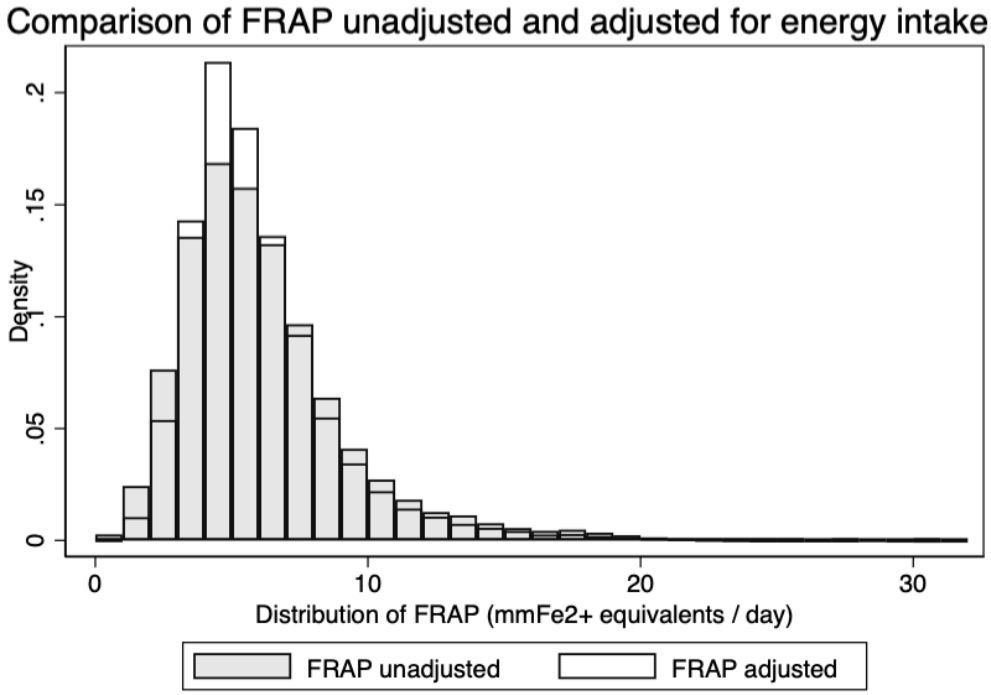

Figure 4.2.2. Energy adjustment of dietary FRAP; FRAP values were obtained from the antioxidant food database developed by Carlsen et al.; energy adjustment applied on the data from the Swedish Women's Lifestyle and Health Cohort.

On the other hand, we are interested in adjusting for energy intake to control for confounding, which can result if energy intake is associated with disease risk. Therefore, a variation of the nutrient residual model includes both, the nutrient residual plus the term for total energy intake, which is presented in Model 2 in **Table 4.2.1**. With this model we receive the overall effect of the nutrient on the disease, letting the nutrient residual represent the dietary composition and allowing total energy intake to have its standard biological meaning [94].

4.2.4 Comparison of antioxidant food databases

Because for studies I, III and IV we were interested in using both antioxidant food databases, a comparison of the two databases was done using data from the WLHC.

The antioxidant food database by Carlsen et al., covering 73 out of 81 items in the FFQ was more complete compared to the antioxidant database by Pellegrini et al., which covered only 39 items of the FFQ. When looking at single food items, the available FRAP values from both databases were very similar, with exception of grain products and few vegetable items showing relevantly higher FRAP values in the database of Pellegrini et al. compared to the values reported in Carlsen et al. Moreover, the items found only in the database of Carlsen et al. corresponded mainly to dairy and meat products, which had very low FRAP values.

Nevertheless, the overall baseline distribution of FRAP comparing both databases was similar (**Table 4.2.2**).

Table 4.2.2. Baseline distribution of dietary NEAC[a] in the Swedish Women's Lifestyle and Health Cohort – Comparison of antioxidant food databases.

Overall distribution of dietary NEAC					
	Mean	Median	Min.	Max	
Carlsen et al.	6.01	5.39	0.44	31.08	
Pellegrini et al.	6.92	6.34	0.28	38.94	
Quintiles of dietary NEAC					
	Q1	Q2	Q3	Q4	Q5
Carlsen et al.	0.4-4.0	4.1-4.9	4.9-6.0	6.0-7.6	7.7-31.1
Pellegrini et al.	0.3-4.5	4.6-5.7	5.8-7.00	7.0-8.8	8.9-38.9

[a] NEAC was assessed by a validated food frequency questionnaire and estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+} equivalents/day

Table 4.2.3. Contribution of food items to total dietary FRAP in the Swedish Women's Lifestyle and Health Cohort- Comparison of antioxidant food databases.

Contribution of food groups to total dietary FRAP, in %							
Database	Vegetables	Fruits	Grains	Alcohol	Chocolate	Tea	Dairy products
Carlsen et al.	12	27	10	2	8	27	3
Pellegrini et al.	11	24	16	3	10	25	0

When comparing the two databases, FRAP values showed a good correlation (Pearson's correlation r=0.7). Further, the percentages of food items contributing most to total dietary NEAC were comparable between the databases, although the contribution of grains was somewhat higher in the database by Pellegrini et al. (**Table 4.2.3**). When comparing classification of the subjects into quintiles of dietary NEAC, there is an overall good accordance between the two databases (**Table 4.2.4**).

Table 4.2.4. Comparison of classification of subjects into quintiles of dietary NEAC, using the FRAP assay and data from the Swedish Women's Lifestyle and Health Cohort.

Exposure classification of Subjects – Categorization into quintiles of NEAC						
Overall number of subjects per quintile	Pellegrini et al.					
Carlsen et al.	Q1	Q2	Q3	Q4	Q5	Total
Q1	7,547	1,315	247	58	10	9,177
Q2	1,618	5,693	1,522	304	39	9,176
Q3	11	2,160	5,371	1,491	144	9,177
Q4	1	7	2,036	6,025	1,107	9,176
Q5	0	1	1	1,298	7,876	9,176
Total	9,177	9,176	9,177	9,176	9,176	45,882

4.3 Assessment of outcome variables

Sweden has a long history of national registers, which allows register-based research with data kept by governmental agencies or other organizations. Access to most of the registers is available through Statistics Sweden (SCB) and the Swedish National Board of Health and Welfare (Socialstyrelsen).

4.3.1 The Swedish Personal Identity Number

The Swedish personal identity number (PIN) is a unique identifier, which has been assigned to every individual with a permanent residence in Sweden [96]. The PIN was introduced in 1947 and allows linkages of individuals to National Demographic and Health Registries. Back then the PIN consisted of the date of birth and a three-digit number. In 1967 a check digit was added, which should verify that date of birth and the three-digit number are correct. All notifications with updates are stored in a notification database held by Statistics Sweden [96]. With permission from the National Board of Health and Welfare, using the PIN we

were able to link the two Swedish cohorts to the Swedish Inpatient Register, the Cancer Register and the Causes of Death Register.

4.3.2 Swedish National Registers

The Swedish National Inpatient Register (IPR) was initiated in 1964 and holds information on dates of hospital admission and discharge, information on the primary diagnosis and secondary diagnoses. Full coverage was reached in 1987. Diagnoses are coded according to the Swedish version of the international coding of disease (ICD) system (ICD-7: 1964-1968; ICD-8: 1969-1986; ICD-9: 1987-1996; ICD-10: 1996 onwards) [97]. The Swedish IPR has a high validity for most of the diagnoses and positive predictive values (PPV) are generally ranging from 85-95%. Because of the long follow-up the register is particularly suitable for large-scale population-based studies [97].

The Swedish Cancer Register is widely used for research purposes and for monitoring cancer incidence and survival. The register covers the whole Swedish population since 1958. It provides detailed information on patient data, medical data (site of tumor, histology, date of diagnosis) and follow-up data. There are some discrepancies between cases reported to the Swedish Cancer Register only and the Swedish IPR only, which leads to some underreporting of cases in the cancer register (3.7% in 1998). Underreporting was shown to be cancer site specific and increased with age of the patients. Nevertheless, the overall completeness is high and comparable to other high quality registers in Northern Europe, and the underreporting might not have major impact on epidemiological studies [98].

The Cause of Death Register is used for official statistics and for medical research. The register provides information on all deaths in Sweden since 1952, including the underlying cause of death. To make international comparisons between causes of death possible the register uses the international version of the ICD coding system developed by the World Health Organisation. The register is virtually complete, with 96% of individuals having a recorded specific underlying cause of death, and can be regarded as a largely complete and high quality source for research purposes [99].

In addition we have used the Statistic Sweden's Register of the Total Population and Population Changes (TPR) to obtain dates of relocation and emigration. The register contains information on date of birth, death, name change, marital status, family relationships and migration and immigration. By using the PIN for all residents staying at least one year in Sweden, the TPR can be used for medical research purposes. The TPR is constantly updated

by the Tax Agency. The long history and the high completeness of the TPR make the register to an important resource for epidemiological studies [100].

4.3.3 Identification of cases and follow-up

Cases of MI, stroke and HF were identified by linking the Swedish PIN through records to the Swedish Inpatient and Causes of Death Register. We used the Swedish international classification of disease (ICD) system, which was adapted from the WHO ICD classification system, using ICD -7 (1964-68), 8 (1968-87) and 9 (1987-97), to identify cases before the beginning of follow-up, and ICD-10 (1997 onwards) to identify cases during follow-up (**Table 4.3.1**) [97]. For heart failure (Study IV) we further considered only hospitalization or mortality of HF as a primary diagnosis as the definition of outcome [101]. For study II, we further distinguished between fatal and non-fatal cases by classifying a subjects as a fatal case if a) MI was identified through the Death Register and b) a subjects died within a 28 day period after MI diagnosis [65].

Table 4.3.1. Classification of cases of selected outcome variables based on the Swedish International Coding of Diseases ICD-7, 8, 9 and 10.

Outcome	**ICD7**	**ICD8**	**ICD9**	**ICD10**
Heart Failure	434.1 434.2 440.99 441.99	427.00 427.10 428.99 402.99	402 428.A 428.AA 428.B 428.BA 428.BB 428.X	I110 I50
Myocardial infarction	420.10 420.17	410	410	I21
Stroke	330.10-330.99 331.00-331.02 331.09 331.999 332.00 332.11 332.12 332.19 332.29 334.09	430 431.00 431.08 431.09 431.90 431.98 431.99 432-434 436	430 431 432.X 433 434 436	I60 I61 I63.0-163.5 I63.8 I63.9 I64
CVD	400-459	390-459	390-459	I00-I99

Follow-up of the Swedish Women's and Lifestyle Cohort

The WLHC was followed from the date of enrolment, which varied between 1991 and 1992, until the time of the event of interest occurred, subjects emigrated or died, or the end of the study period, which was December 31, 2012, whichever occurred first.

Exclusion criteria for the WLHC are presented in the flow-chart in **Figure 4.3.1**. For all three studies (I, III, IV) we excluded women who had been diagnosed with any cardiovascular disease (n=1,983) and cancer (except non-melanoma skin cancer) (n=716), before the start of follow-up, using ICD-8 and 9. Further, to reduce exposure misclassification induced by self-reported information in the FFQ, we excluded women who reported implausible values for energy intake (± 3SDs from the $\log_e$-transformed mean energy intake) (n=502) and who had left the whole or the main section of the FFQ blank (n=493), leaving us a final cohort of 45,882 women.

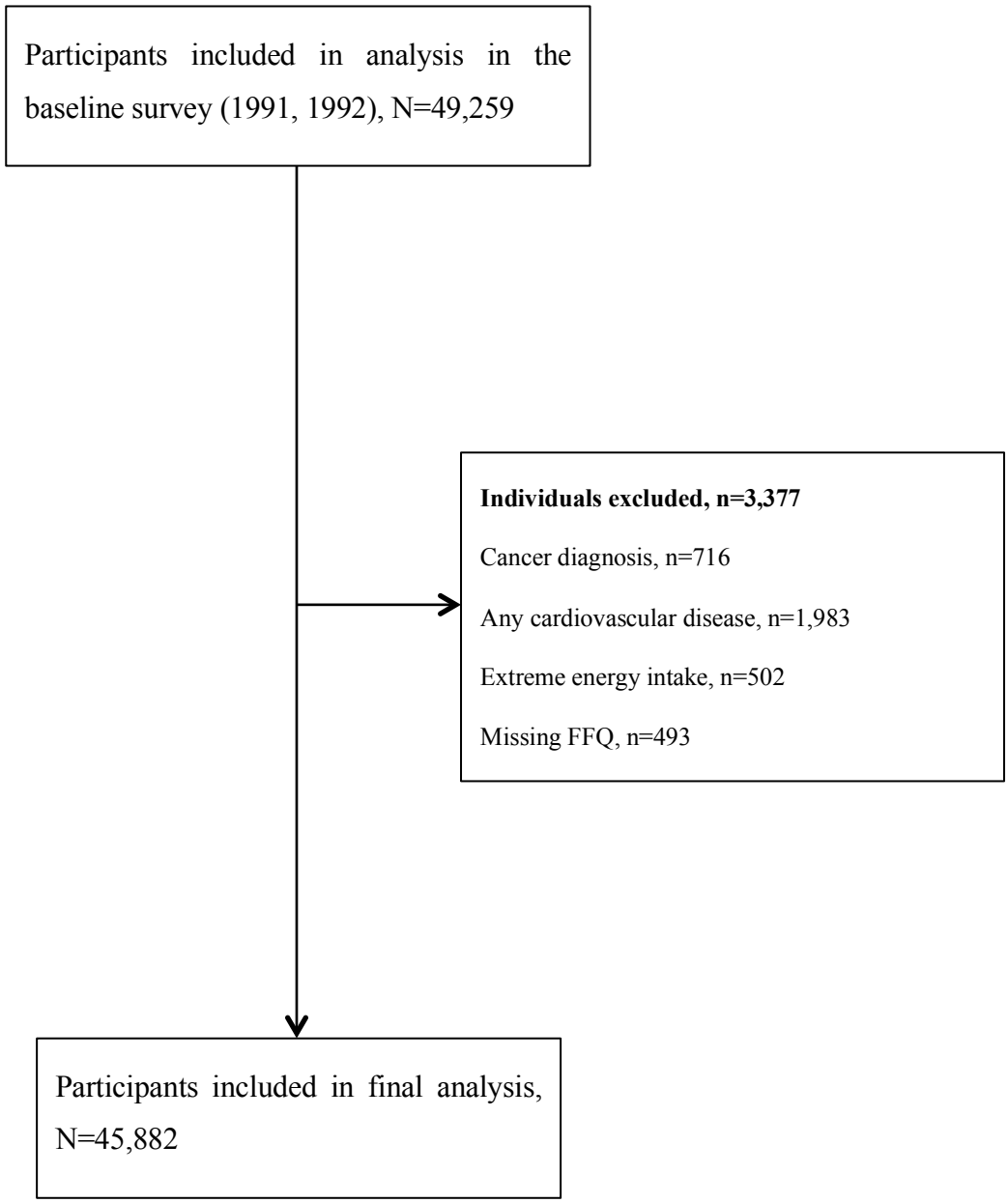

Figure 4.3.1. Flow-chart, Source population -The Swedish Women's Lifestyle and Health Cohort - Exclusion Criteria.

Follow-up Swedish National March Cohort

The SNMC was followed from October 1, 1997, until the time of the event of interest occurred, subjects emigrated or died, or the end of the study period, which was December 31, 2010, whichever occurred first.

Exclusion criteria for the cohort are presented in the flow-chart in **Figure 4.3.2**. For this study we excluded participants who provided an incorrect personal number (n=11), who were below the age of 18 years (n=1,740) and who emigrated (n=466) or died (n=9) before the beginning of the follow-up. We further excluded participants with a previous diagnosis of cancer (except non-melanoma skin cancer) (n= 2,673), or any cardiovascular disease (n= 4,733) before the beginning of follow-up, using ICD-8 and 9. Further, to reduce exposure misclassification induced by self-reported information in the FFQ we excluded subjects who reported extreme energy intakes (± 3 SDs of the mean value of the $\log_e$-transformed energy intake) at baseline (n=470) [75], leaving 34,543 subjects for the final analysis.

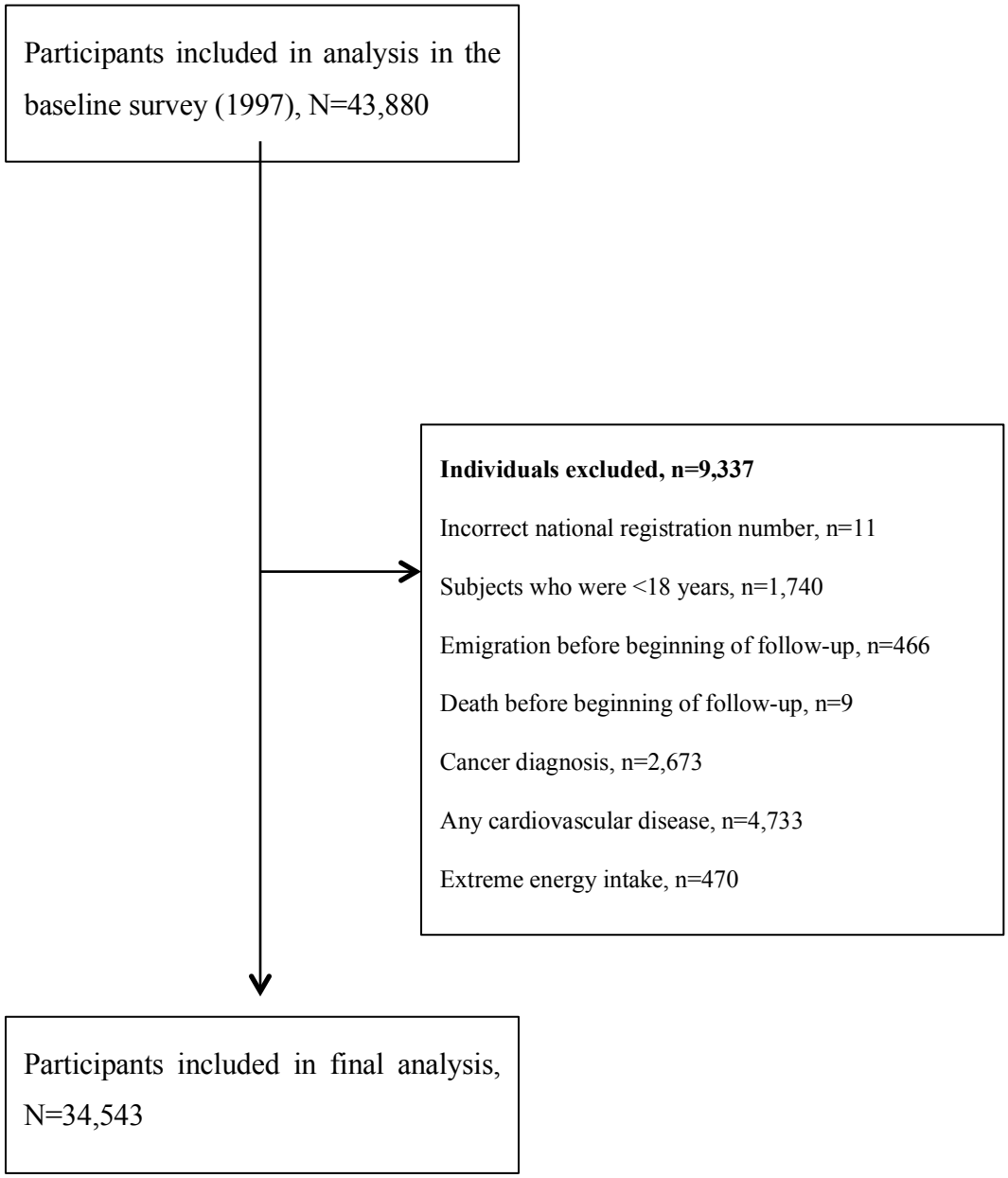

Figure 4.3.2. Flow-chart, Source population -The Swedish National March Cohort - Exclusion Criteria.

4.4 Assessment of covariates

The baseline questionnaires used in the WLHC and the SNMC covered detailed information on lifestyle factors and medical history. In this section the assessment of important covariates and their use for the specific studies will be described.

Swedish Women's Lifestyle and Health Cohort

The baseline questionnaire in the WLHC was used for studies I, III and IV and covered information on anthropometric measures, educational level, physical activity, alcohol consumption, smoking status, use of multivitamin supplements, as well as self-reported history of diabetes and hypertension.

To compute body mass index (BMI) the reported weight was divided by the height squared (kg/m^2). To assess educational level women were asked to report their total number of years of education, which was categorized into "≤10 years", "11-13 years" and ">13 years". Physical activity was assessed by asking women to report their level of physical activity, which ranged from one (=very low) to five (=very high). For the purpose of the study levels one and two were combined into "low" and levels four and five into "high" activity. To estimate daily alcohol consumption, the reported glasses of beer, wine and liquor were converted into total grams of alcohol per day. The alcohol variable was further categorized into levels of low (<5g/day), medium (5-25g/day) and high (>25g/day) consumption. Smoking habits were assessed by asking women, if they have ever smoked regularly, and by assessing the number of cigarettes smoked during different age periods. This information was used to categorize women into "no", "former" and "current" smokers. Information on multivitamin supplement use was collected within the question covering general use of vitamin and mineral supplements. Women were asked to report whether they had ever consume supplements, with further information requested for the type of supplements used and the number of supplements consumed per week and over the year. As the reported information on duration and frequency was not very accurate, women were categorized in "users" and "non-users" only. History of diabetes and hypertension was assessed by asking women to report, if they have ever been treated for one of these health conditions (yes, no).

Swedish National March Cohort

The baseline questionnaire in the SNMC was used for study II and provided information on anthropometric measures, educational level, alcohol consumption, smoking status, use of vitamin supplements, aspirin use, as well as self-reported history of diabetes, lipid disturbance, and hypertension.

Body mass index (BMI) was computed as weight divided by height squared (kg/m^2). Educational level was assessed by asking subject to report their type of education (e.g. elementary school, secondary school, university/high school), which was then categorized into levels of "<13 years" and "≥13 years" of education. Physical activity during a typical day was assessed through a validated Energy Expenditure Questionnaire [102, 103], that asked subjects to report how much time they have spent in different activities during an ordinary 24h-day. This allowed us to estimate total MET-hours per day (METh/day), where MET stands for metabolic energy turnover with one MET corresponding to an energy expenditure of 1 kcal/kg body weight per hour [104]. Alcohol consumption was assessed through the FFQ by converting the reported glasses of beer, wine and liquor into total grams of alcohol consumed per month (g/month). Smoking status was assessed by asking subjects to report if they have ever smoked cigarettes daily for more than six months, and by asking to report the number of cigarettes smoked during different age periods. This information was then used to categorize the variable into "never", "former", "current (≤ 15 cigarettes/day) and "current (>15 cigarettes/day)" smokers. Information on vitamin supplement use was collected by asking subjects for the type of vitamins used, the overall duration of consumption (never, less than one year, one to five years, more than five years) and whether consumption differed during the year, e.g. due to seasonality. Because of inaccurate information on the type of vitamin used and the duration, we created a general variable for vitamin and mineral supplement use (users, non-users). Aspirin use was assessed by asking subjects to report their average use of aspirin during the previous year, previous five years ago and previous ten years ago, with consumption frequencies ranging from 0, 1-5, 6-10, 11-20, 21-30 and more than 30 tablets per month. This information was used to categorize subjects into "users", if subjects had ever consumed aspiring during any of the time periods, and "non-users" otherwise. History of diabetes, hypertension and lipid disturbance was assessed by asking subjects to report whether they have ever been treated for one of these health conditions by a doctor (yes, no).

5 STATISTICAL METHODS

In this paragraph the theoretical concepts of survival analysis and the Cox model will be introduced. Further, some other statistical methods and their theoretical background will be explained. The final section describes the implementation of the statistical methods in studies I to IV.

5.1 Concepts of Survival Analysis

When conducting prospective cohort studies we are not only interested in the rate of the event of interest over time, but also in the time until the event occurs, which we refer to as survival time [105]. There are some unique features coming with this kind of data, which make survival analysis important.

First, as indicated before, we are interested in two quantities, the event or failure, and the time to the event T. Since time T is always positive and commonly skewed, we will generally find large numbers of events either at the beginning or at the end of follow-up [105].

Second, by the end of follow-up not everyone will have had the event of interest and the true time to the event is unknown for those individuals. The phenomenon of unknown survival time for a subset of the study population is called censoring. Censoring can either occur because a) a subject did not experienced the event of interest by the end of follow-up, b) a subject dropped out of the study during follow-up, or c) a subject experienced a different event, which makes further follow-up impossible. Censored survival times will always underestimate the true time to event. If the event of interest, which we assume is yet to occur, is beyond the end of follow-up, we talk about right censoring [105].

Survival data can be described in terms of two probabilities. The survival probability, also called the survivor function S(t), is the probability of a subject to survive from the time origin to a specific future time t. The hazard function h(t) is the instantaneous event rate for the disease for an individual who has already survived until time t, with T representing the random variable for an individual's survival time [105, 106].

$$h(t) = \lim_{\Delta \to 0} \frac{\Pr\,(t + \Delta t > T > t \,\mathrm{I}\, \mathrm{T} > \mathrm{t})}{\Delta t} \tag{5.1}$$

There is a clear relationship between S(t) and h(t), as denoted by the following calculus formula [105]:

$$h(t) = -\frac{d}{dt}[logS(t)] \tag{5.2}$$

Given the equation we can see, that if either S(t) or h(t) is known, we can automatically determine the other.

There are specific methods, which can be used to estimate the above quantities, of which two are presented below.

Kaplan Meier estimator

The Kaplan-Meier estimator, also known as product limit method, is a nonparametric method, that allows us to estimate survival probability from observed survival times, both censored and uncensored. Based on the assumption, that the events occur independently, we can estimate the cumulative survival probability by multiplying the probability of surviving from one interval to the next based on the following calculus formula.

$$S(t_j) = S(t_{j-1})(1 - \frac{d_j}{n_j}) \tag{5.3}$$

As in observational studies we further want to condition the results on other predictors, the Kaplan-Meier method is generally used as a preliminary analysis [105].

Cox Proportional Hazards Regression

The Cox Proportional Hazards regression model is a multivariate approach used to assess survival data given a set of covariates [107]. The relationship between the event incidence, which is expressed by the hazard function h(t), and the covariates is described in formula (5.4), where h_0 is the baseline hazard and corresponds to the value of the hazard if all covariates X_i were equal to zero.

$$h(t) = h_0(t)e^{\Sigma_{i=1}^{p}\beta_i X_i} \tag{5.4}$$

The Cox model can be seen as a multiple linear regression of the logarithm of the hazard given the variables X_i, with $h_0(t)$ being the intercept that varies over time t. The model is

semi-parametric because the baseline hazard can take any form while the covariates enter the model linearly [107].

A key assumption of the model is the Proportional Hazards (PH) assumption [108]. It states that the hazard of the event in any group is a constant multiple of the hazard in the other group. This further implies, that given the covariates are time-independent, the hazard ratio (HR) between two or more groups is constant over time, which can be shown by the following calculus.

$$HR = \frac{h_0(t)e^{\sum_{i=1}^{p}\beta_i X_i^*}}{h_0(t)e^{\sum_{i=1}^{p}\beta_i X_i}} = e^{\sum_{i=1}^{p}\beta_i(X_i^* - X_i)} \tag{5.6}$$

It is important to verify that the PH assumption holds, which can be done in different ways. One subjective approach is to test graphically, whether the plotted hazard curves of each group are proportional and if they cross each other. Unfortunately, this method does not incorporate the effect of covariates, and therefore convergent or divergent lines could either be due to real lack of proportionality or because an important confounder was not considered [108]. Another approach is to use a formal statistical test, e.g. Schoenfeld residuals. If the PH assumption holds for a particular covariate, this implies that the Schoenfeld residuals of that covariate are not related to the survival time [109]. If the PH assumption is violated, one can alternatively fit a stratified model, where the covariate displaying non-proportionality is fitted without constraints of proportionality. As such a covariate will not have any estimated effect in a stratified model, this approach should only be used for covariates, which are not of primary interest for the analysis [108].

5.2 Other methodology

Dose-response analysis

To investigate the dose-response relationship between exposure and outcome restricted cubic splines (RCS) can be used. RCS is a powerful tool used to characterize the relationship between a continuous exposure and an outcome by visually and statistically testing the assumption of a linear association [110]. Splines are flexible functions defined by piecewise polynomials and junctions between them, so called knots, while restricting the fitted function to be linear before the first and after the last knot [111]. Placing knots at fixed percentiles of

the marginal distribution of the exposure ensures a uniform allocation of points within each interval and further prevents outliers from influencing the placement of the knots [112].

Effect modification

It is not unusual for an effect of one exposure and outcome to depend on the presence or absence of another exposure, which leads to the concept of interaction [25, 113]. The definition of interaction is equivalent to the definition of effect-measure modification, if no bias is present [25]. Interaction can be measured on two scales, the additive and multiplicative scale.

A natural way to assess, whether interaction occurs is to measure to what extent the effect of the two factors together exceed the effect of each factor considered individually, which is assessing interaction on the additive scale. Interaction on additive scale can be tested using the relative excess risk due to interaction (RERI) [114]. RERI measures the risk that is additional to the expected risk, by calculating the observed risk RR_{11}, when both exposures are present, minus the expected risk $(RR_{10} + RR_{01} - 1)$, when both exposures occurre separately. A RERI>1 indicates positive additive interaction and a RERI<1 indicates negative additive interaction.

$$\begin{aligned} RERI_{RR} &= RR_{11} - (RR_{10} + RR_{01} - 1) \\ RERI_{RR} &= RR_{11} - RR_{10} - RR_{01} + 1 \end{aligned} \tag{5.7}$$

Other measures for additive interaction are the synergy index (S) and the proportion attributable to interaction (AP). S measures to which extent the effect of both exposures together exceed 1, and if this is greater than the sum of to which extent the effect of both exposures separately exceed 1.

$$S = \frac{RR_{11} - 1}{(RR_{10} - 1) + (RR_{01} - 1)} \tag{5.8}$$

AP is a derivative measure of RERI and measures the proportion of the risk in the group exposed to both factors that is due to the interaction itself.

$$AP = \frac{RR_{11} - RR_{10} - RR_{01} + 1}{RR_{11}} \tag{5.9}$$

Based on a simulation study conducted by Li & Chambless RERI should be the measure of choice when additive interaction is assessed in a proportional hazards model, which is why RERI was chosen as the primary indicator for interaction in studies I-IV [115].

On the other hand, we can measure whether the effect of both factors together exceed the product of the two factors considered separately, which is referred to assessing interaction on the multiplicative scale. A measure of multiplicative interaction is given by the following calculus [113]. If $RR_{11}/(RR_{10}RR_{01})>1$, the interaction is said to be positive, and negative if $RR_{11}/(RR_{10}RR_{01})<1$ [113].

$$\frac{RR_{11}}{RR_{10}RR_{01}} = \frac{p_{11}p_{00}}{p_{10}p_{01}} \tag{5.10}$$

It should be noted, that sometimes we have a positive interaction on the additive but a negative interaction on the multiplicative scale, or vice versa. In addition, interaction may be present on one scale but not the other. The choice of the scale depends mainly on the motivation of the study. Nevertheless, it has been suggested to always report both additive and multiplicative measures of interaction [113].

Multiple imputation

Missing data is a challenge in epidemiological research. Especially observational studies are prone to missing data, because even in the scarce setting where there is no missing information for the exposure and outcome, covariates necessary to adjust for confounding might be affected by missingness. Data can be either a) missing completely at random (MCAR), where missingness does not depend on any other observed variable, or b) missing at random (MAR), where the probability that a given subset of variables is observed depends only on the values of observed variables, or c) missing not at random (MNAR), where missingness depends on unobserved variables. For epidemiologic studies it is reasonable to assume that data is MAR. One way to address missingness under the assumption of MAR, is conducting a complete-case analysis, i.e. excluding all individuals with missing values. However, this approach may yield biased results [116]. Alternatively, it is possible to impute missing values by conducting multiple imputations, e.g. based on chained equations, which

accommodates arbitrary missing-value patterns [117]. This will create a pre-specified number of imputed datasets, which can be used to assess a single estimate by pooling estimates from the imputed datasets using Rubin's formula for standard errors [118].

5.3 Data analyses

The statistical methods used for all studies were fairly similar and the main analyses are therefore presented together for studies I, III and IV, which were conducted within the WLHC, and separately for study II, which was conducted within the SNMC.

5.3.1 Studies I, III & IV

To analyze the effect of dietary NEAC on the risk of MI, stroke and HF in women, dietary NEAC was categorized into quintiles in study I and III and tertiles in study IV. The advantage of categorizing the exposure into quantiles in nutritional epidemiology includes the reduction of potential effects of outlying or extreme values. In addition, by using categories we do not make assumptions about the dose-response relationship, which allows us to visualize if the association is most likely following a linear or non-linear trend [92]. A drawback, however, is, that if the relationship is truly linear, categorization of the exposure will likely lead to a loss of power [119].

We then fitted Cox proportional hazards regression models with age as underlying time scale [120] to estimate hazard ratios (HRs) with 95% confidence intervals (CIs). The exposure variable was implanted both, as a continuous variable and as a categorical variable, by considering the first quantile of dietary NEAC as the reference category for the latter.

Potential confounders were selected based on subject matter knowledge. In addition, directed acyclic graphs (DAGs), which offer a systematic representation of causal relationships [121], were used to help evaluate the choice of potential confounders and to reduce bias caused by over adjustment [25]. To do so, we used the web-based application DAGitty[1], which helps to draw and analyze DAGs (the DAG is presented in **Figure A.1** in the Appendix). The models were adjusted for the following confounders,: education (≤10 years, 11-13 years, >13 years), Body Mass Index (BMI, kg/m^2), smoking (no, former, current), physical activity level (low, medium, high), total alcohol intake (<5g/day, 5-25g/d, >25g/day), total energy intake (kcal/day), multivitamin supplements use (yes, no), self-reported hypertension (yes, no) and

[1] The web-based application DAGitty can be found online under http://dagitty.net.

diabetes (yes, no). Because coffee was not included in the assessment of dietary NEAC, the model was further adjusted for coffee intake (0 cups/day, >0-4 cups/day, >4 cups/day) to control for potential bias.

In study IV the main model was additionally adjusted for incident myocardial infarction diagnosed before heart failure as a time varying covariate. This was done by splitting single time-span records into episodes before and after myocardial infarction using the Stata command stsplit [122] and by implementing the new time varying covariate in the model.

The proportional hazards assumption was tested using Schoenfeld residuals. If the statistical test indicated a violation of the assumption, stratified analyses were conducted for the corresponding variables. To assess linear trends we created new categorical variables based on the median of each NEAC quantile. This variable was then implemented in both, unadjusted and adjusted models as a continuous variable.

To further assess the dose-response relationship we used restricted cubic spines and placed knots at the 5^{th}, 35^{th}, 65^{th} and 95^{th} percentile of the distribution of dietary NEAC for studies I and III, whereas for study IV three knots were chosen at the 10^{th}, 50^{th} and 90^{th} percentile.

Subgroup analyses were conducted for the variables BMI (<25, ≥25 kg/m2), smoking (non-current, current) and vitamin supplement use (yes, no). Potential effect modification was investigated on both multiplicative and additive scale. On multiplicative scale the likelihood ratio test was used to formally test, whether interaction was present. To do so we ran the Cox model once by adding a cross-product interaction term between dietary NEAC and the variable of interest in the model, and once without. After that the likelihood ratio test was used to assess which model fitted the data better. On the additive scale, the RERI was used to assess whether interaction was present.

As a main sensitivity analysis we further assessed whether the estimates stayed robust when using the antioxidant food database developed by Pellegrini et al. to assess dietary NEAC by the FRAP and TRAP assay. Percentages of missing values were 6.4% for diabetes, 4.5% for physical activity, 4.4% for hypertension, 3.7% for BMI, 1.9% for education and 0.3% for smoking. To see whether the observed association was affected by missing values. To do so we repeated the main analyses after imputing missing values for the confounders using multiple imputation based on chained equations. Fifteen imputed datasets were created to assess a single estimate by pooling HRs using Rubin's formula for standard errors [118]. In addition, analyses were repeated after excluding cases occurring during the first two years of follow-up to address potential reverse causation.

5.3.2 Study II

To study the association between dietary NEAC and the risk of overall, fatal and non-fatal MI, dietary NEAC was categorized into sex-specific quartiles. We then fitted Cox proportional hazards regression models with age as underlying time scale [120] to estimate hazards ratios (HRs) with 95% confidence intervals (CIs). The exposure variable was implemented as a continuous variable and as a categorical variable, with the first quantile of dietary NEAC as the reference category.

We adjusted the models for the following potential confounders: sex, BMI (kg/m2, continuous), total physical activity (METh/day, continuous), alcohol consumption (g/month, continuous), smoking status (never, former, current [≤15 cigarettes/day, >15 cigarettes/day]), level of education (<13 years, ≥13 years), history of diabetes (yes/no), aspirin use (yes/no), coffee consumption (cups/day; 0, 1–2, 3–4, ≥5), self-reported lipid disturbance (yes/no), self-reported hypertension (yes/no), vitamin supplement use (yes/no) and energy intake (kcal/day, continuous).

Again, we created a new categorical variable based on the median of each NEAC quartile to assess the linear trend and implemented it in both, unadjusted and adjusted models, as a continuous variable. The proportional hazards assumption was tested using Schoenfeld residuals. If the statistical test indicated a violation of the assumption, stratified analyses were conducted for the corresponding variables.

To further assess the dose-response relationship we used restricted cubic splines and placed knots at the median of each quartile of dietary NEAC.

We conducted subgroup analyses for the variables BMI (<25, ≥25 kg/m2), smoking (non-current, current) and vitamin supplement use (yes, no), sex (male, female), alcohol consumption (no/yes) and age at baseline (<60/≥60 years). Potential effect modification was investigated on both multiplicative and the additive scale.

We further conducted main sensitivity analyses by repeating the analyses using dietary NEAC assessed through the TRAP assay as main exposure. Percentages of missing values were equal to 10.3% for total physical activity, 7.7% for smoking status, 4.3% for BMI, 4.2% for history of hyperlipoproteinemia, 4.0% for aspirin use, 3.8% for history of diabetes, 3.2% for history of hypertension, 1.7% for coffee consumption, 1.3% for vitamin supplement use, 1.0% for level of education and 0.1% for alcohol drinking. To assess whether the observed association was affected by missing values the analyses were repeated after imputing missing values for the confounders using multiple imputation based on chained equations. To do so

fifteen imputed datasets were created to assess a single estimate by pooling HRs using Rubin's formula for standard errors [118]. Further, analyses were repeated after excluding cases occurring during the first three years of follow-up to investigate potential effects of reverse causation.

6 MAIN FINDINGS

In this paragraph the baseline characteristics for the Swedish Women's Lifestyle and Health Cohort and the Swedish National March Cohort are presented, followed by the main findings for studies I to IV.

All previously published material was reproduced with the permission from the publisher.

6.1 Baseline Characteristics - WLHC

Baseline characteristics of the Swedish Women's Lifestyle and Health Cohort are presented in **Table 6.1**. The characteristics are presented for the total population and based on quintiles of dietary NEAC.

Mean age at enrolment was 39.5 years (SD=5.7). In the WLHC 30.8% of the women had completed more than 13 years of education, 25% were overweight or obese at baseline and 21% reported to smoke at the time of enrollment to the study. Further, 81% of the women reported to be physically active and 75% consumed less than 5g of alcohol per day. 8.7% of the women reported to be post-menopausal and 3% to receive hormone replacement therapy (HRT).

When comparing women across quintiles of dietary NEAC, women in the highest quintile were better educated and reported to be non-smokers. They further tended to consume higher amounts of fruits, vegetables, tea as well as multivitamin, vitamin and mineral supplements. Women in the lowest quintile of dietary NEAC had a higher BMI, consumed more coffee and alcohol and further reported more hypertension and diabetes.

Table 6.1. Baseline characteristics of the study population presented for the overall population and across quintiles of dietary NEAC –Swedish Women's Lifestyle and Health Cohort, 1991-2012.

Quintiles of dietary NEAC[a]	Total sample	Q1	Q2	Q3	Q4	Q5
	6.0 (2.8)	≤4.0	4.0-4.9	4.9-6.0	6.0-7.6	7.6-31.1
Characteristics	n = 45,882	n = 9,177	n = 9,176	n = 9,177	n = 9,176	n = 9,176
Age at baseline, mean, SD	39,5 (5.7)	39.4 (5.7)	39.6 (5.8)	39.7 (5.7)	39.4 (5.8)	39.4 (5.7)
Education, %						
> 13 years	30.8	20.1	26.1	30.5	34.9	45.1
BMI, mean, SD	23.4 (3.7)	23.5 (3.9)	23.6 (3.8)	23.5 (3.6)	23.3 (3.5)	23.2 (3.5)
Energy intake (kcal/day), mean, SD	1580 (480)	1560 (510)	1610 (470)	1605 (480)	1590 (445)	1535 (475)
Physical activity, %						
low	14.2	18.6	14.1	14.1	13.2	14.3
medium	57.0	62.7	62.6	59.3	58.9	54.9
high	24.3	18.7	23.2	26.7	27.8	30.7
Smoking, %						
current	21.2	30.9	23.5	20.6	16.8	14.6
Saturated fat (g/day), mean, SD	23.6 (5.4)	25.3 (6.1)	23.8 (5.1)	23.1 (5.0)	23.0 (5.0)	22.7 (5.5)
Monounsaturated fat (g/day), mean, SD	17.5 (3.1)	18.1 (3.3)	17.6 (2.9)	17.3 (2.9)	17.3 (2.9)	16.9 (3.3)
Polyunsaturated fat (g/day), mean, SD	7.4 (1.7)	7.4 (1.9)	7.5 (1.7)	7.4 (1.6)	7.3 (1.6)	7.3 (1.7)
Alcohol intake, %						
<5 g/day	75.1	74.7	75.3	75.0	75.2	75.4
5-25 g/day	24.5	24.4	24.3	24.7	24.6	24.3

> 25 g/day	0.4	0.9	0.4	0.3	0.3	0.3
Fruits (servings/day), mean, SD	1.9 (1.1)	1.0 (0.7)	1.8 (0.8)	2.0 (1.0)	2.1 (1.1)	2.4 (1.3)
Vegetables (servings/day), mean, SD	2.3 (1.0)	2.0 (0.9)	2.3 (0.9)	2.4 (1.0)	2.5 (1.0)	2.5 (1.2)
Grains (servings/day), mean, SD	5.4 (2.3)	5.3 (2.6)	5.5 (2.2)	5.4 (2.3)	5.3 (2.2)	5.2 (2.2)
Tea (cups/day), mean, SD	0.5 (0.7)	0.1 (0.1)	0.1 (0.2)	0.3 (0.4)	0.6 (0.4)	1.5 (0.8)
Coffee (cups/day), mean, SD	3.7 (2.5)	4.6 (2.7)	4.2 (2.3)	3.9 (2.4)	3.3 (2.2)	2.5 (2.2)
Multivitamin supplement use (yes), %	14.9	10.9	14.2	14.1	16.6	18.8
Vitamin and mineral supplement use (yes), %	41.5	36.9	41.4	42.8	43.8	46.0
Menopause (yes), %	8.7	9.4	9.4	9.4	8.6	8.3
HRT (yes), %	3.0	3.0	3.2	3.3	3.0	3.1
Hypertension (yes), %	8.8	9.6	9.8	9.1	8.7	8.6
Diabetes (yes), %	1.2	1.5	1.3	1.1	1.3	1.3

[a] NEAC was assessed by a validated food frequency questionnaire and estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+} equivalents/day

6.2 Baseline Characteristics - SNMC

Baseline characteristics of the Swedish National March Cohort are presented in **Table 6.2**. The characteristics are presented for the total population and based on quartiles of dietary NEAC.

Mean age at enrolment was 49.4 years (SD=15.8) and 65.7% of the cohort were women. 29.4% of the subjects had completed more than 13 years of education. Mean BMI of the study population was 24.5 kg/m^2 and 8.1% reported to smoke at that time. In addition, 9.9% reported to have hypertension, 2% to have diabetes and 2.3% to be treated for lipid disturbance. Moreover, 49.6% of subjects consumed vitamin and mineral supplements and 17.3% reported to consume aspirin on more than six times per month. Women had a slightly higher absolute daily dietary NEAC at baseline compared to men.

When comparing subjects across quartiles of dietary baseline NEAC, subjects in the fourth quartile were slightly older, were better educated, tended to be non-smokers and consumed more vitamin and mineral supplements. In addition, they drank less coffee, but had a higher consumption of tea and alcohol. They also tended to report to be treated to hypertension and diabetes to a slightly higher degree.

Table 6.2. Baseline characteristics of study participants presented for the overall population and across quartiles of dietary NEAC[a], Swedish National March Cohort, 1997-2010.

Sex-specific quartiles of dietary NEAC	Total sample	Q1	Q2	Q3	Q4
Females		<7.0	7.0-9.2	9.3-12.1	12.2-46.5
Males		<6.3	6.3-8.1	8.2-10.6	10.7-42.9
Characteristics	n=34,543	n=8,636	n=8,636	n=8,636	n=8,635
Age (years) mean (SD)	49.4 (15.8)	47.6 (16.1)	49.0 (16.0)	49.8 (15.8)	51.4 (15.2)
Female, %	65.7	65.7	65.7	65.7	65.7
BMI (kg/m^2), mean (SD)	24.5 (3.5)	24.8 (3.6)	24.6 (3.5)	24.4 (3.5)	24.2 (3.3)
Total physical activity (METh/day),	39.5 (12.8)	39.7 (13.6)	39.8 (12.9)	39.5 (12.4)	39.1 (12.0)
Alcohol (grams/month), mean (SD)	306.8 (619.5)	263.9 (411.4)	304.3 (494.9)	317.8 (486.8)	341.1 (938.5)
Energy intake (kcal/day), mean (SD)	2040 (540)	1990 (540)	2090 (535)	205 (540)	1995 (530)
Tea (cups/day), mean, SD	1.0 (1.2)	0.0 (0.2)	0.4 (0.5)	0.9 (0.7)	2.3 (1.4)
Coffee (cups/day), mean, SD	2.9 (1.8)	3.4 (1.9)	3.1 (1.8)	2.8 (1.7)	2.2 (1.7)
Diabetes (yes), %	2.0	1.9	2.0	2.1	2.1
Lipid disturbance (yes), %	2.3	2.2	2.3	2.5	2.3
Aspirin use (yes), %	17.8	18.8	18.1	17.8	19.8
High blood pressure (yes), %	9.9	9.0	9.9	9.8	10.8
Vitamin and supplement use (yes), %	49.6	44.8	48.4	51.2	53.9
Education (≥13 years), %	29.4	22.0	26.1	30.8	38.8
Smoking, %					
Never	65.0	61.7	64.5	65.9	67.9
Former	26.9	26.7	26.9	27.2	26.8
Current (≤15 cigarettes/day)	7.6	10.7	8.1	6.6	5.0
Current (>15 cigarettes/day)	0.5	0.9	0.5	0.4	0.4

[a] NEAC estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+}

6.3 Study I – Dietary NEAC and myocardial infarction in women

During a mean follow-up time of 20.3 years we observed 657 cases of incident myocardial infarction, with a mean age at first diagnosis of 56.6 years.

After adjusting for potential confounders a significant 28% lower risk for subjects in the fourth (HR: 0.72; 95% CI: 0.55-0.95) and a 40% lower risk for subjects in the fifth quintile of dietary NEAC (HR: 0.60, 95% CI: 0.45-0.81), compared to subjects in the first, was found, with a significant trend (p-value < 0.001) (**Table 6.3.1**). In addition, when implementing dietary NEAC as a continuous variable, the risk of MI decreased by each unit increment of dietary NEAC by 5% (HR: 0.95; 95% CI: 0.91-0.98). When further investigating the dose-response relationship no deviation from a linear association was detected (p-value for non-linearity > 0.05) (**Fig. 6.3.1**), confirming a linear association between dietary NEAC and the risk of MI.

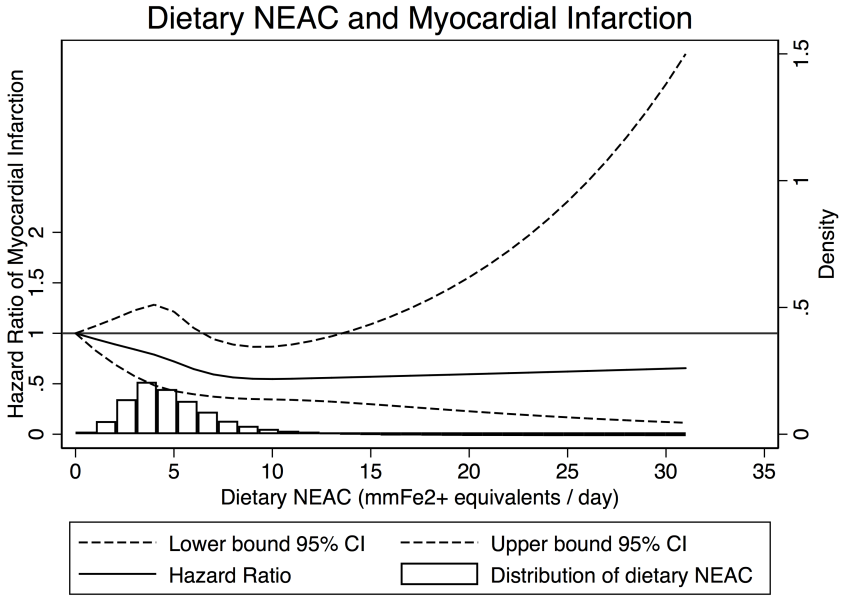

Figure 6.3.1. Multivariable-adjusted restricted cubic spline curve for the relation between dietary Non Enzymatic Antioxidant Capacity (NEAC), measured in mmol $Fe2^{+}$ equivalents per day, and risk of myocardial infarction. Adjustments were made for age, education (≤10 years, 11-13 years, >13 years), BMI (kg/m^2), smoking (no, former, current), physical activity (low, medium, high), total alcohol intake (<5g/day, 5-25g/d, >25g/day), total energy intake (kcal/day), multivitamin supplements use (yes, no), hypertension (yes, no), diabetes (yes, no).

Subgroup analyses suggested a somewhat stronger effect in overweight and obese subjects, non-current smokers and vitamin and mineral supplement users. However, no evidence for effect modification was found on the multiplicative (all p-values > 0.05) or the additive scale (all p-values for RERI > 0.05).

When further conducting sensitivity analyses, estimates remained essentially the same after using FRAP (HR Q5 vs. Q1: 0.71, 95% CI: 0.54-0.94) and TRAP (HR Q5 vs. Q1: 0.70, 95% CI: 0.52-0.93) from the secondary database to assess dietary NEAC and after imputing missing values (HR Q5 vs. Q1: 0.63; 95% CI: 0.49-0.82).

Table 6.3.1. Hazard ratios (HRs) with 95% confidence intervals (95% CIs) for dietary NEAC[a] in relation to the risk of myocardial infarction - Swedish Women's Lifestyle and Health Cohort, 1991-2012.

Quintiles of dietary Non Enzymatic Antioxidant Capacity (NEAC)[a]						
	Q1	Q2	Q3	Q4	Q5	*p* for trend
No. of cases	196	145	126	105	85	
Person-years	184,911.75	186,266.62	186,421.78	186,489.92	185,456.82	
Incidence rate (per 100,000 person-years)	106.00	77.85	67.59	56.30	45.83	
Age-adjusted HR (95% CI)	1.0 (reference)	0.71 (0.58-0.89)	0.62 (0.49-0.77)	0.52 (0.41-0.66)	0.43 (0.33-0.55)	<0.001
Multivariable HR (95% CI)[b]	1.0 (reference)	0.90 (0.71-1.15)	0.79 (0.61-1.02)	0.72 (0.55-0.95)	0.60 (0.45-0.81)	<0.001

[a] NEAC assessed by a validated food frequency questionnaire, estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+} equivalents / day, adjusted for energy intake

[b] Adjusted for age, education (≤10 years, 11-13 years, > 13 years), BMI (kg/m^2), smoking (no, former, current), physical activity (low, medium, high), total alcohol intake (<5g/day, 5-25g/d, >25g/day), total energy intake (kcal/day), multivitamin supplements use (yes, no), hypertension (yes, no), diabetes (yes, no)

6.4 Study II – Dietary NEAC and myocardial infarction in men and women

In this study the association between dietary NEAC and the risk of myocardial infarction was investigated in men and women who participated in the Swedish National March Cohort. During a mean follow-up time of 12.7 years a total of 1,142 cases of incident MI, of which 205 were fatal and 937 non-fatal, were identified. Foods contributing most to total dietary NEAC in the SNMC were tea (26%), fruits (21%), vegetables (20%), grains (9%) and chocolate (5%).

When investigating total myocardial infarction the proportional hazards assumption was violated for sex in the multivariate model and therefore a stratified Cox model was fitted. Results are presented in **Table 6.4.1**. After adjusting the model for potential confounders the risk among subjects in the fourth quartile of dietary NEAC was 23% lower (HR: 0.77, 95% CI: 0.61-0.96; *p* for trend = 0.008) compared to subjects in the first. When running Cox regression for men and women separately, a similar HR was found for men (HR: 0.78; 95% CI: 0.59-1.04; *p* for trend = 0.056) and women (HR: 0.73; 95% CI: 0.49-1.08; *p* for trend = 0.040), however, the association was not significant in men. When further investigating the dose-response relationship, no deviation from a linear association was seen for overall (*p*-value for non-linearity = 0.56) (**Figure 6.4.1**).

When investigating fatal and non-fatal MI separately, an inverse association was found only with non-fatal MI in the fully adjusted model with lowest risk in subjects in the highest quartile compared to the first (HR: 0.72; 95% CI: 0.56-0.92; *p* for trend < 0.05). We did not detect any departure from linearity after conducting spline analyses for non-fatal MI (*p*-value for non-linearity > 0.05). When running the analyses for men and women separately, an inverse association was detected in women for non-fatal MI (HR Q4 vs. Q1: 0.62, 95% CI: 0.41-0.96; *p* for trend < 0.05), whereas for men the effect remained again short of significance (HR: 0.75; 95% CI: 0.55-1.02; *p* for trend > 0.05). Findings for non-fatal and fatal MI are presented in **Table 6.4.2** and **Table 6.4.3**.

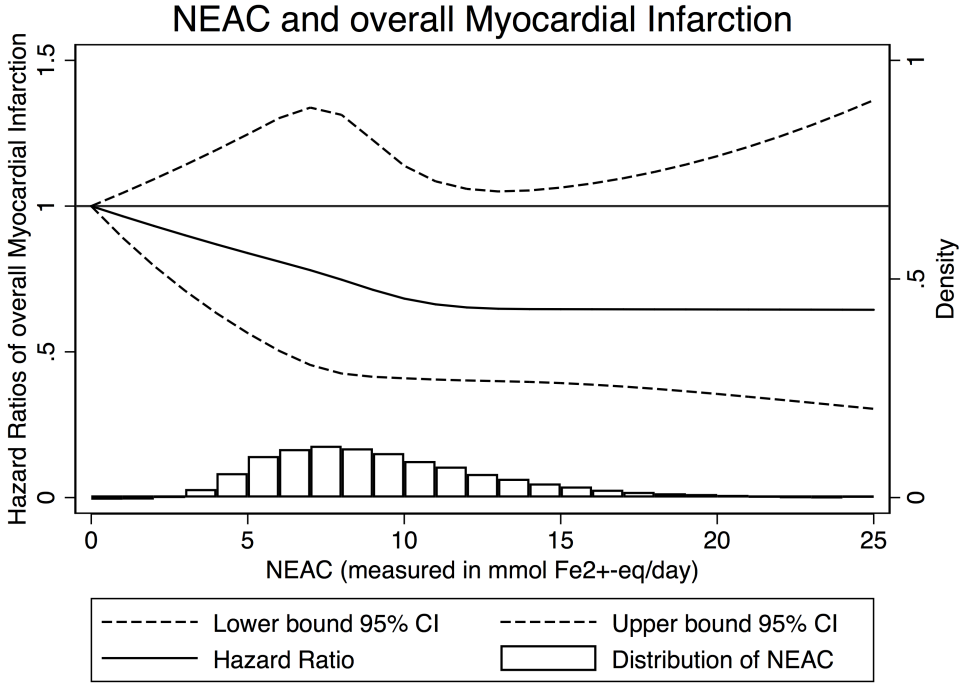

Figure 6.4.1. Multivariable-adjusted restricted cubic spline curve for the relation between dietary Non Enzymatic Antioxidant Capacity (NEAC), measured in mmol Fe2+ equivalents per day, and the risk of overall myocardial infarction. Adjustments were made for age, (sex), education (<13 years, ≥13 years), smoking (no, former, current [≤15 cigarettes/day; >15 cigarettes/day]), total alcohol intake (g/month), coffee (0, 1-2, 3-4, ≥5 cups/day), diabetes (yes, no), BMI (kg/m^2), total energy intake (kcal/day), vitamin and mineral supplements use (yes, no), total physical activity (METh/day), lipid disturbance (yes, no), high blood pressure (yes, no), aspirin use (yes, no).

Further, subgroup analyses suggested a slightly stronger association in non-obese subjects, in never smokers, in alcohol drinkers, in subjects below the age of 60 and in supplement non-users. However, no evidence was found for effect modification on the multiplicative (all *p*-values > 0.05) or on the additive scale (*p*-values for all RERI > 0.05).

In addition, after conducting sensitivity analyses by using the TRAP assay for the assessment of dietary NEAC, by imputing missing values, estimates remained essentially the same for all outcomes.

Table 6.4.1. Hazard ratios (HRs) with 95% confidence intervals (95% CIs) for dietary Non Enzymatic Antioxidant Capacity (NEAC)[a] in relation to the risk of overall myocardial infarction, Swedish National March Cohort, 1997-2010.

	Quartiles of dietary Non Enzymatic Antioxidant Capacity (NEAC)[a]				
	Q1	Q2	Q3	Q4	*p* for trend
No. of cases[b]	292	301	271	278	
Person-years	109 819.07	109 687.52	109 767.58	109 304.98	
Incidence rates[c] (per 100'000 person-years)	532.46	490.41	415.58	428.95	
Total sample					
Age-adjusted HR (95% CI)	1.0 (reference)	0.92 (0.79-1.09)	0.80 (0.68-0.94)	0.75 (0.63-0.88)	0.000
Multivariable HR (95% CI)[d]	1.0 (reference)	0.99 (0.80-1.22)	0.80 (0.65-1.00)	0.77 (0.61-0.96)	0.008
Women					
No. of cases	122	120	108	106	
Age-adjusted HR (95% CI)	1.0 (reference)	0.91 (0.80-1.17)	0.81 (0.62-1.04)	0.74 (0.57-0.96)	0.019
Multivariable HR (95% CI)[d]	1.0 (reference)	1.09 (0.78-1.54)	0.82 (0.57-1.18)	0.73 (0.49-1.08)	0.040
Men					

No. of cases	170	181	163	172	
Age-adjusted HR (95% CI)	1.0 (reference)	0.92 (0.75-1.14)	0.75 (0.60-0.93)	0.69 (0.56-0.86)	0.000
Multivariable HR (95% CI)[d]	1.0 (reference)	0.95 (0.73-1.24)	0.78 (0.59-1.02)	0.78 (0.59-1.04)	0.056

[a] NEAC was assessed by a validated food frequency questionnaire and estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+} equivalents/day

[b] Numbers refer to observations included in the age- and sex-adjusted models

[c] Age adjusted incident rates

[d] Adjusted for age, education (<13 years, ≥13 years), smoking (no, former, current [≤15 cigarettes/day; >15 cigarettes/day]), total alcohol intake (g/month), coffee (0, 1-2, 3-4, ≥5 cups/day), diabetes (yes, no), BMI (kg/m^2), total energy intake (kcal/day), vitamin and mineral supplements use (yes, no), total physical activity (METh/day), lipid disturbance (yes, no), high blood pressure (yes, no), aspirin use (yes, no); (stratified for sex in the overall study population)

Table 6.4.2. Hazard ratios (HRs) with 95% confidence intervals (95% CIs) for dietary Non Enzymatic Antioxidant Capacity (NEAC)[a] in relation to the risk of non-fatal myocardial infarction, Swedish National March Cohort, 1997-2010.

	Quartiles of dietary Non Enzymatic Antioxidant Capacity (NEAC)[a]				
	Q1	Q2	Q3	Q4	*p* for trend
No. of cases[b]	249	241	221	226	
Person-years	109 819.07	109 687.52	109 767.58	109 304.98	
Incidence rates[c] (per 100'000 person-years)	435.82	366.09	322.79	309.28	
Total sample					
Age-adjusted HR (95% CI)	1.0 (reference)	0.87 (0.73-1.04)	0.76 (0.64-0.92)	0.72 (0.60-0.86)	0.000
Multivariable HR (95% CI)[d]	1.0 (reference)	0.94 (0.75-1.19)	0.77 (0.61-0.98)	0.72 (0.56-0.92)	0.004
Women					
No. of cases	109	104	82	89	
Age-adjusted HR (95% CI)	1.0 (reference)	0.88 (0.67-1.15)	0.69 (0.52-0.92)	0.70 (0.53-0.93)	0.008
Multivariable HR (95% CI)[d]	1.0 (reference)	1.06 (0.74-1.51)	0.68 (0.46-1.02)	0.62 (0.41-0.96)	0.008

Men

No. of cases	140	137	139	137	
Age-adjusted HR (95% CI)	1.0 (reference)	0.85 (0.67-1.08)	0.77 (0.61-0.98)	0.68 (0.53-0.86)	0.001
Multivariable HR (95% CI)[d]	1.0 (reference)	0.88 (0.65-1.18)	0.80 (0.59-1.07)	0.75 (0.55-1.02)	0.065

[a] NEAC was assessed by a validated food frequency questionnaire and estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+} equivalents/day

[b] Numbers refer to observations included in the age- and sex-adjusted models

[c] Age adjusted incident rates

[d] Adjusted for age, education (<13 years, ≥13 years), smoking (no, former, current [≤15 cigarettes/day; >15 cigarettes/day]), total alcohol intake (g/month), coffee (0, 1-2, 3-4, ≥5 cups/day), diabetes (yes, no), BMI (kg/m^2), total energy intake (kcal/day), vitamin and mineral supplements use (yes, no), total physical activity (METh/day), lipid disturbance (yes, no), high blood pressure (yes, no), aspirin use (yes, no), (stratified for sex)

Table 6.4.3. Hazard ratios (HRs) with 95% confidence intervals (95% CIs) for dietary Non Enzymatic Antioxidant Capacity (NEAC)[a] in relation to the risk of fatal myocardial infarction, Swedish National March Cohort, 1997-2010.

	Quartiles of dietary Non Enzymatic Antioxidant Capacity (NEAC)[a]				
	Q1	Q2	Q3	Q4	*p* for trend
No. of cases[b]	43	60	50	52	
Person-years	109 819.07	109 687.52	109 767.58	109 304.98	
Incidence rate[c] (per 100'000 person-years)	96.64	124.32	92.80	120.67	
Total sample					
Age-adjusted HR (95% CI)	1.0 (reference)	1.23 (0.83-1.83)	0.99 (0.66-1.49)	0.91 (0.61-1.37)	0.342
Multivariable HR (95% CI)[d]	1.0 (reference)	1.27 (0.77-2.11)	0.95 (0.55-1.63)	1.10 (0.63-1.91)	0.980
Women					
No. of cases	13	16	26	17	
Age-adjusted HR (95% CI)	1.0 (reference)	1.13 (0.55-2.36)	1.79 (0.92-3.49)	1.07 (0.52-2.21)	0.800
Multivariable HR (95% CI)[d]	1.0 (reference)	1.68 (0.62-4.55)	2.29 (0.87-6.03)	1.90 (0.68-5.28)	0.265
Men					

No. of cases	30	44	24	35	
Age-adjusted HR (95% CI)	1.0 (reference)	1.24 (0.78-1.98)	0.62 (0.36-1.06)	0.77 (0.47-1.25)	0.077
Multivariable HR (95% CI)[d]	1.0 (reference)	1.17 (0.65-2.12)	0.57 (0.29-1.15)	0.86 (0.44-1.66)	0.353

[a] NEAC was assessed by a validated food frequency questionnaire and estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+} equivalents/day

[b] Numbers refer to observations included in the age- and sex-adjusted models

[c] Age adjusted incident rates

[d] Adjusted for age, education (<13 years, ≥13 years), smoking (no, former, current [≤15 cigarettes/day; >15 cigarettes/day]), total alcohol intake (g/month), coffee (0, 1-2, 3-4, ≥5 cups/day), diabetes (yes, no), BMI (kg/m^2), total energy intake (kcal/day), vitamin and mineral supplements use (yes, no), total physical activity (METh/day), lipid disturbance (yes, no), high blood pressure (yes, no), aspirin use (yes, no), (stratified for sex)

6.5 Study III – Dietary NEAC and stroke in women

During a mean follow-up time of 20.2 years 881 incidence cases of stroke, of which 521 were ischemic, 297 hemorrhagic and, 63 unspecified stroke, were found, with a mean age at first stroke of 55.9 years.

In the age adjusted model an inverse association between daily dietary NEAC and overall stroke with lowest hazard rates in the fifth quintile compared to the first was found (HR: 0.60, 95% CI: 0.49-0.73, *p* for trend < 0.001). However, after inclusion of potential confounders, no clear pattern was seen anymore (**Figure 6.5.1**). Findings were similar when conducting separate analyses for ischemic and hemorrhagic stroke (*p* values for trend > 0.05) (**Figure 6.5.1**). The dose-response relationship showed no deviation from a linear association (all *p*-value for non-linearity > 0.05).

When further investigating potential effect modification, we did not detect any difference in the effect of dietary NEAC on the risk of stroke in subgroups of BMI, smoking and vitamin supplement use on any of the stroke outcomes on the multiplicative (all *p*-values > 0.05) or additive scale (all *p*-values for RERI > 0.05).

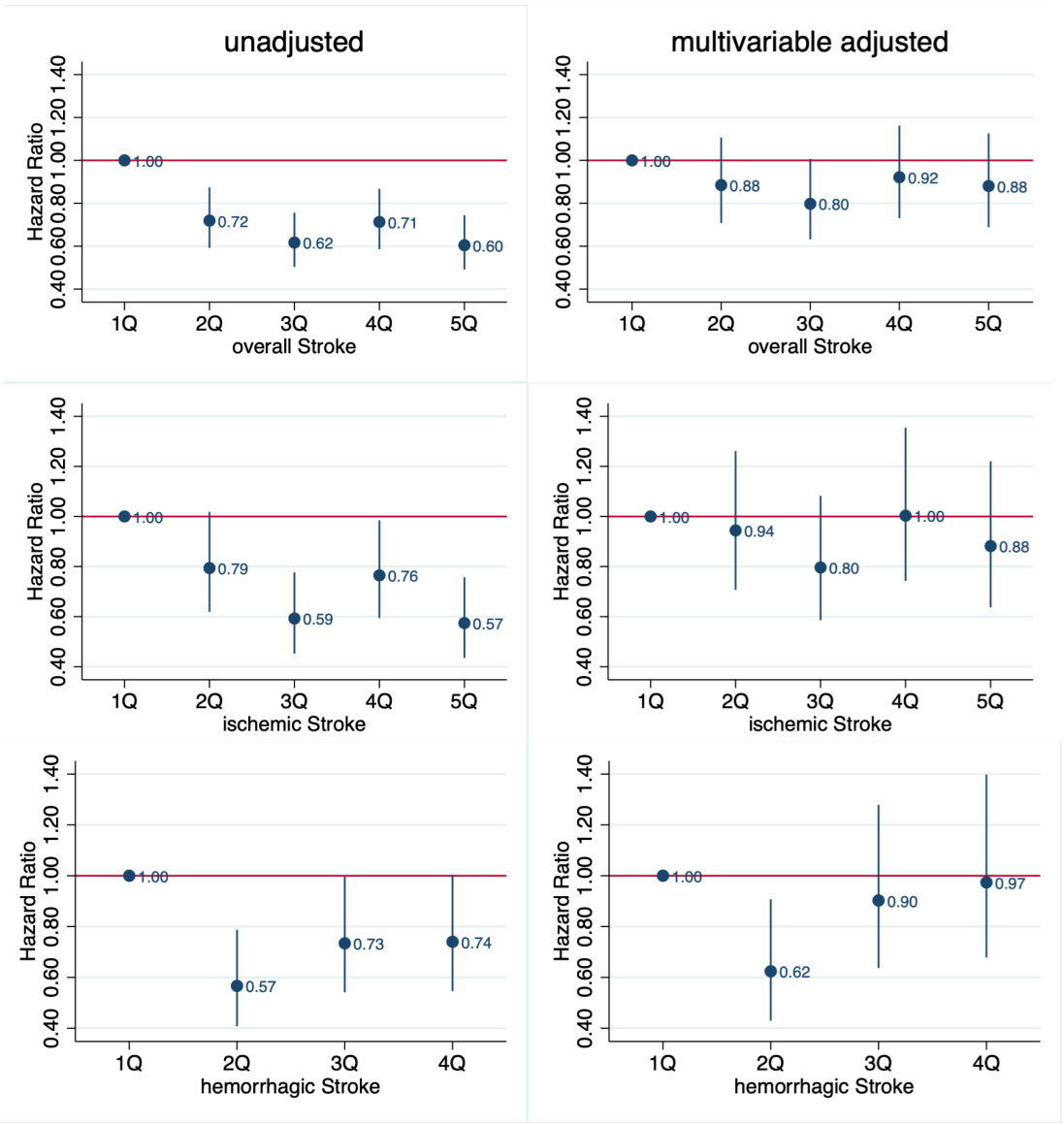

Figure 6.5.1. Hazard ratios (HRs) with 95% confidence intervals (95% CIs) for dietary Non Enzymatic Antioxidant Capacity (NEAC), measured in mmol $Fe2^{+}$ equivalents per day, in relation to the risk of total, ischemic and hemorrhagic stroke, Swedish Women's Lifestyle and Health Cohort, 1991-2012. The left column represents unadjusted estimates. The right column represents multivariable adjusted estimates. Models were adjusted for age (time scale), education (≤10 years, 11-13 years, > 13 years), BMI (kg/m^2), smoking (no, former, current), physical activity (low, medium, high), total alcohol intake (<5g/day, 5-25g/d, >25g/day), total energy intake (kcal/day), multivitamin supplement use (yes, no), high blood pressure (yes, no), diabetes (yes, no), coffee intake (0, 0-4, >4 cups/day).

6.6 Study IV – Dietary NEAC and heart failure in women

After a mean follow-up time of 20.3 years, 245 incidence cases of heart failure were detected, with a mean age at first diagnosis of 58.3 years.

An inverse association between daily dietary NEAC and incidence or mortality of heart failure as the primary cause was found, with a significant 27% lower risk in subjects in the highest compared to the lowest tertile (HR: 0.63; 95% CI: 0.43-0.93; *p* for trend < 0.05) (**Table 6.6.1**). When including dietary NEAC as a continuous variable in the model each unit increment of dietary NEAC was associated with a 6% risk reduction of heart failure (HR: 0.94; 95% CI: 0.88-1.00). When further investigating the dose-response relationship using restricted cubic splines, the linear association was further confirmed, showing no deviation from linearity (*p*-value for non-linearity > 0.05) (**Figure 6.6.1**).

Subgroup analyses suggested a slightly stronger effect in overweight subjects, current smokers and vitamin supplement users. However, no evidence for effect modification on the multiplicative (all *p*-values above > 0.05) or the additive scale (all *p*-values for RERI > 0.05) was found.

After conducting main sensitivity analyses using the antioxidant database developed by Pellegrini et al. to assess dietary NEAC, the effect was slightly attenuated for FRAP (Q3 vs Q1: HR: 0.67, 95% CI: 0.46-0.98, *p* for trend < 0.05) and for TRAP (Q3 vs Q1: HR: 0.70, 95% CI: 0.48-1.02, *p* for trend > 0.05). Similarly, after imputing missing values estimates remained essentially the same (HR Q3 vs. Q1: 0.67; 95% CI: 0.48-0.93; *p* for trend < 0.05).

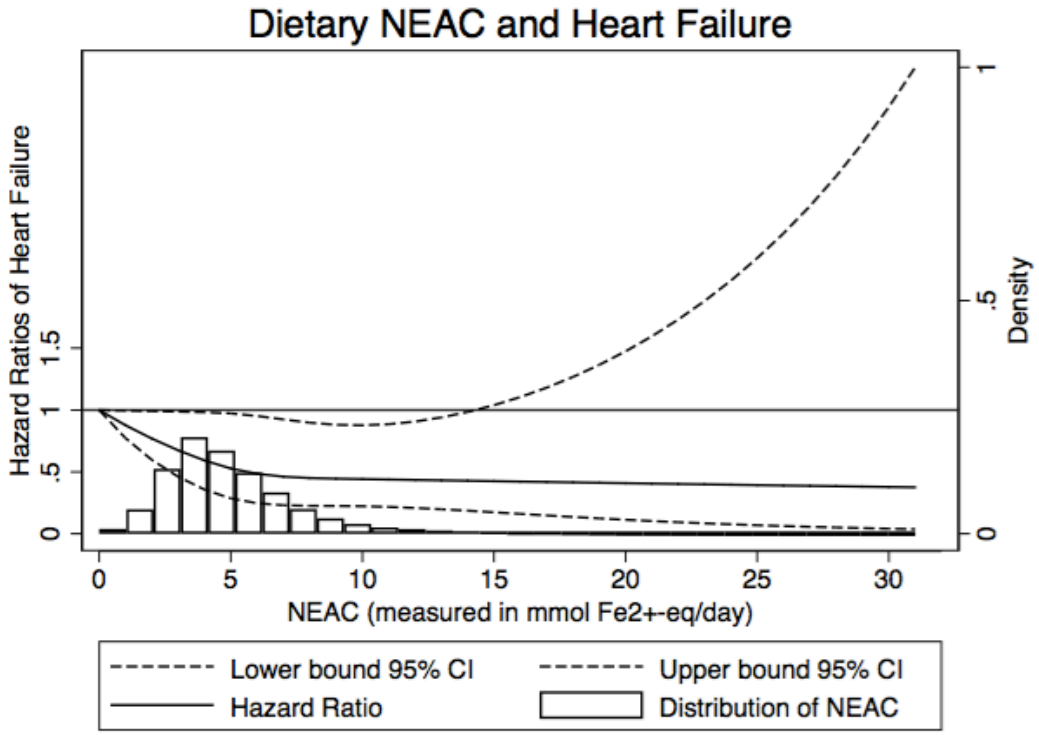

Figure 6.6.1. Multivariable-adjusted restricted cubic spline curve for the relation between dietary Non Enzymatic Antioxidant Capacity (NEAC), measured in mmol $Fe2^{+}$ equivalents per day, and the risk of heart failure. Adjustments were made for age, education (≤10 years, 11-13 years, > 13 years), BMI (kg/m2), smoking (no, former, current), physical activity (low, medium, high), total alcohol intake (<5g/day, 5-25g/d, >25g/day), total energy intake (kcal/day), multivitamin supplement use (yes, no); Stratified for high blood pressure (yes, no), diabetes (yes, no), incident myocardial infarction as time-varying covariate (yes, no).

Table 6.6.1. Hazards ratios (HRs) with 95% confidence intervals (95% CIs) for dietary NEAC[a] in relation to the risk of heart failure, Swedish Women's Lifestyle and Health Cohort, 1991-2012.

Tertiles of dietary Non Enzymatic Antioxidant Capacity (NEAC)[a]				
	Q1	Q2	Q3	*p* for trend
No. of cases	117	79	58	
Person-years	310,297.45	311,998.96	310,034.90	
Incidence rates (per 100,000 person-years)	37.71	25.32	18.71	
Age-adjusted HR (95% CI)	1.0 (reference)	0.66 (0.49-0.87)	0.50 (0.36-0.68)	<0.001
Multivariable HR (95% CI)[b]	1.0 (reference)	0.78 (0.56-1.09)	0.63 (0.43-0.93)	<0.05

[a] NEAC was assessed by a validated food frequency questionnaire, estimated through the ferric reducing antioxidant power (FRAP) assay, expressed in mmol Fe^{2+} equivalents / day, adjusted for energy intake.

[b]Model was adjusted for age (time scale), education (≤10 years, 11-13 years, > 13 years), BMI (kg/m2), smoking (no, former, current), physical activity (low, medium, high), total alcohol intake (<5g/day, 5-25g/d, >25g/day), total energy intake (kcal/day), multivitamin supplement use (yes, no); Stratified for high blood pressure (yes, no), diabetes (yes, no) and incident myocardial infarction as time-varying covariate (yes, no) because of violation of the proportion hazards assumption.

7 DISCUSSION

This paragraph is separated into four subsections. First, the interpretations of the main findings for studies I to IV are highlighted. Second, some methodological considerations and the generalizability of the findings will be discussed. In the last two sections some final remarks on antioxidants and NEAC and the final conclusions will be presented.

7.1 Interpretation of the findings

7.1.1 Study I

In Study I a higher baseline dietary NEAC was associated with a lower risk of MI in women aged 30-49 years participating in the Swedish Women's Lifestyle and Health Cohort and there was evidence for a linear relationship. The effect of dietary NEAC on MI risk was slightly stronger in obese, non-current smokers and vitamin and mineral supplement users. However, there was no evidence for effect modification when using statistical formal tests. Nevertheless, the statistical power might be limited due to the smaller number of subjects and cases in each subgroup. After conducting the analyses using NEAC values based on FRAP and TRAP from the secondary database the estimates remained similar, showing a significant inverse association.

These findings are in line with the results reported from two previous observational studies. In the Swedish Mammography Cohort study, which consists of 38,984 women aged 49 to 83 years, higher dietary NEAC, measured through the ORAC assay, was inversely associated with the risk of myocardial infarction [75]. In an Italian case-control study involving 760 cases and 682 controls a higher dietary NEAC, measured through the FRAP, TRAP and TEAC assay, was inversely related with the risk of non-fatal acute myocardial infarction for each assay separately [76]. Further, the effect was substantially stronger in women compared to men.

To conclude, this study confirms previous evidence of higher dietary NEAC to be associated with a lower risk of MI and gives novel insight in the effect in young and middle aged women.

7.1.2 Study II

Study II, which was conducted within men and women participating in the Swedish National March Cohort, provided evidence that a higher dietary baseline NEAC to be associated with a lower risk of MI, specifically total and non-fatal MI, whereas no association was found for

fatal MI. Further, the effect might be stronger in women compared to men. When conducting further subgroup analyses a slightly stronger association was found in non-obese subjects, in never smokers, in alcohol drinkers, in subjects who were below the age of 60 years and in subjects who did not consume any vitamin and mineral supplements. However, based on formal tests there was no significant evidence for effect modification. Again, statistical power might be an issue because of the smaller number of subjects and cases within each subgroup.

Findings from Study II are further in line with findings reported in Study I. In addition, they confirm the results reported from the Swedish Mammography Cohort [75] and the Italian case-control study [76]. In the latter, the inverse association was stronger in women, which was confirmed by our study, but contrary the effect was stronger in subjects aged 60 years or older [76].

Overall, this study supports previous findings and provides novel evidence that a higher dietary NEAC lowers the risk of total and non-fatal MI, but not fatal MI. It further highlights potential differences of the effect between men and women.

7.1.3 Study III

The aim of Study III was to investigate the association between dietary NEAC and the risk of stroke. However, after adjusting the model for potential confounders, the initially significant inverse association diminished and we found no relationship between dietary NEAC and the risk of overall, ischemic or hemorrhagic stroke in young and middle-aged women in the Swedish Women's Lifestyle and Health Cohort. This implies that the effect might have strongly been confounded by other lifestyle factors, such as smoking and physical activity.

Findings from previous studies investigating the effect of dietary NEAC and stroke risk are contradicting. In the Italian segment of the European Prospective Investigation into Cancer and Nutrition study (EPIC) comprising 41,620 men and women aged 35 to 65 years, higher dietary NEAC, measured through the TEAC assay, was associated with a reduced risk of ischemic stroke, but not overall or hemorrhagic stroke [78]. In the Swedish Mammography Cohort, dietary NEAC measured through the ORAC assay was associated with a lower risk of overall stroke among CVD-free women and with a lower risk of hemorrhagic stroke among women with a CVD history. No association was found for ischemic stroke [77]. Recently, in 34,555 subjects aged 18 to 94 years participating in the Swedish National March Cohort, a higher dietary NEAC measured through the FRAP and TRAP assay was related to a lower risk of overall and ischemic stroke among women, but not men [79]. In contrast to these findings, the Rotterdam Study found no effect between dietary NEAC measured

through the FRAP assay on stroke risk among 5,395 men and women above the age of 55 years [80].

There are different possible explanations to these discrepant findings. First, different databases and chemical assays have been used to assess dietary NEAC [65]. However, when we repeated our analyses using FRAP and TRAP values from the secondary database [67, 68], we did not find any difference in the effect of dietary NEAC on stroke risk. Therefore, the use of different databases or assays is an unlikely explanation for the discrepant findings.

Differences in the findings could be further related to the different contribution of food items to dietary NEAC. For example, the decision, whether to include coffee in NEAC assessment or not, differed between studies. In the Swedish Mammography Cohort, an absorption rate of 6% was considered for coffee. In the Rotterdam study, coffee was fully included in total dietary NEAC, and as coffee holds a high antioxidant capacity value, coffee had a major impact on total dietary NEAC (90%) in their study. In our study the foods contributing most to dietary NEAC were comparable to the proportions reported in the Swedish National March Cohort [79] and the Swedish Mammography Cohort [77]. Food composition of dietary NEAC is therefore an unlikely explanation for our null findings.

Finally, our cohort was rather young at baseline (30-49 years) and showed a mean age of 55 years at first incidence of stroke, which is fairly low compared to the average age of first stroke reported in other studies within Swedish women (e.g. Hallström et al. 2008 with average age of 80.4 years [123]). It is important to note that risk factors for stroke in women change during lifespan. Whilst in young women stroke is mainly related to pregnancy and hormonal factors, risk factors during mid-life, such as high blood pressure, cholesterol levels and waist circumference, increase to a much steeper rate in women compared to men [124]. Considering this specific set of risk factors for young and middle aged women, developing stroke at younger ages could underlie a different biological mechanism compared to stroke developed at older ages. Therefore, diet could affect the risk among age groups in a different way. In the INTERSTROKE study, a large case-control study conducted among 32 countries, a stronger effect of diet on the risk of stroke was seen in men and women above the age of 55 [125]. However, Rautiainen et al. as well as Colarusso et al. found no effect modification in women by age when comparing subgroups of <65 and ≥ 65 or ≤ 60 and >60 years of age at baseline, respectively. Nevertheless, the effect of diet together with age and changing risk factors in women during lifespan should be considered in future studies.

To summarize, study III found no association between dietary NEAC on the risk of overall, ischemic or hemorrhagic stroke in young and middle-aged women. As evidence from previous studies is still limited, the association between dietary NEAC and the risk of stroke should be further explored in other populations.

7.1.4 Study IV

Study IV found a higher baseline dietary NEAC to be associated with a lower risk of heart failure in young and middle aged women in the Swedish Women's Lifestyle and Health Cohort. The effect was further slightly stronger in overweight subjects, current smokers and vitamin and mineral supplement users, although statistical tests did not confirm any evidence for effect modification. However, again statistical power might be limited due to an even smaller number of cases within each subgroup analyzed, as the total number of incidence cases of HF was low in this cohort (n = 245). Again, the association remained statistically significant when assessing dietary NEAC through the FRAP and TRAP assay from the secondary database.

To our knowledge, only one previous study has evaluated the relation between dietary NEAC and the risk of HF. In the Swedish Mammography Cohort a higher dietary NEAC, measured through the ORAC assay, was inversely associated with HF risk among middle aged and older women, suggesting a linear dose-response relationship. Similarly, in our study there was evidence for a linear association between dietary NEAC and HF risk.

Overall, study IV contributes important evidence to the ongoing discussion on the effect of antioxidants on the risk of heart failure, especially in young and middle aged women.

7.2 Methodological considerations

In this section methodological considerations, strengths and limitation of the approaches and potential sources of bias will be discussed, which could have affected the validity of the findings.

7.2.1 Study design

The main strengths of the studies I to IV were their prospective design, the long follow-up and the large sample size. In addition, the baseline questionnaires used in the WLHC and the SNMC assessed information on lifestyle factors and medical history in a great detail. However, some potential sources of bias have to be considered, of which selection bias can

play an important role. Selection bias occurs due to systematic differences in selecting the study groups, which could cause the relation between exposure and outcome to differ between those who were selected in the study and for all those who should have been eligible for the study. This can affect the internal validity of the study and can contribute to under- or overestimation of the actual effect of the exposure on the outcome [25]. Although prospective cohort studies are less prone to suffer from selection bias, as this kind of bias is related to both, the exposure and the outcome and the outcome has not occurred at baseline, three special cases causing selection bias in prospective cohort studies are self-selection bias, loss-to follow-up and competing risk.

Self-selection bias

Self-selection bias can occur due to self-selection of subjects into a study. If the reason for self-selection is associated with the outcome of interest, this can affect the validity of the study [25]. In study II participants for the cohort were self-selected and baseline information was collected among motivated people during a fund raising event, who were particularly interested in supporting cancer research. Therefore the SNMC could be prone to healthy volunteer bias. In fact, compared to the general Swedish population in 1997, cohort members were less educated, were more overweight and obese but smoked to a lesser degree [86], which is why self-selection bias cannot be ruled out for study II. As for studies I, III and IV women were randomly-selected from the general Swedish population living in the Uppsala Health Care region, self-selection bias does most likely not present an issue for the SWLHC. Nevertheless, the cohort might suffer from a healthy volunteer bias, as only 51% of the invited women returned the questionnaire, indicating that women with a serious illness or disease might less likely have participated at baseline [82].

Loss to follow-up

Loss of subjects during follow-up prevents the direct measurement of average risks of a disease, because the outcome of lost subjects is unknown [25]. Therefore it is desirable to keep the loss to follow-up as little as possible. In studies I to IV subjects were followed by linking their individual personal numbers to National Health Registers, which guarantees an essentially complete follow-up, therefore limiting any bias caused by loss to follow-up.

Competing risk

Competing events, which are events that prevent us to observe the outcome of interest, can introduce bias when estimating conditional risks [25]. In populations with high frequency of competing events it is important to appropriately account for them, as it could otherwise lead to falls interpretations of the results [126]. For CVDs a potential competing event is death from other causes. If subjects die before having CVD, they will no longer be at risk for having CVD. Depending on the aim of the study, two approaches to deal with competing risk are commonly used. It has been suggested that in etiologic research the cause-specific hazards model for the cause of interest should be used [126]. This is equivalent to implementing the Cox proportional hazards model while treating competing events as censored ones [127]. In predictive research the Fine and Gray model, which links the effect of risk factors directly to the cause-specific cumulative incidence of the competing event using subdistribution hazards ratios, is often more suitable [126]. Studies I to IV were etiologic studies examining the association between exposure and outcome. Therefore, choosing the Cox model by accounting for the competing event with censoring is appropriate. Nevertheless, in study IV an additional competing risk analysis with death as the competing event was conducted using a flexible parametric model estimating cause-specific hazards for each competing event [128], which further confirmed our findings.

7.2.2 Confounding

Confounding plays an essential role in observational studies and occurs when the effect of the exposure of interest is distorted by extraneous factors, i.e. the confounding factors. Not controlling for confounding can affect internal validity of the study and can further lead to over- or underestimation of an effect, or even change the direction of the measured effect [25].

Choice of confounders

The decision for the choice of confounders for the main analyses was made based on prior knowledge from the literature. In addition, we used DAGs to identify minimal adjustments to estimate the total and direct effect of dietary NEAC on the risk of CVDs. Based on the DAG each variable was suggested to be a confounder, as all of them were suggested to be ancestors of the exposure and the outcome. Nevertheless, the direction of the association can be discussed for some of the variables. Hypertension, diabetes and lipid disturbance are

preclinical conditions and could be intermediate variables as well, as dietary NEAC reduces oxidative stress and might therefore have an effect on these conditions. However, based on the design of the study, subjects have been treated for hypertension, diabetes or lipid disturbance before participation in the study. In this case, the preclinical condition appeared before the assessment of the exposure, which makes the assumption of the direction of the association plausible.

Residual confounding

Although we were able to carefully control for several confounders, some residual confounding might still be present [25].

First of all, some confounding factors could not be considered in our studies. For example, we did not have information on family history of CVD, a factor associated with both, the exposure and any cardiovascular outcome. Someone with family history might be more aware to follow a healthy lifestyle, including a healthy diet, to reduce his or her risk to develop CVD later in life, which could have biased our estimates. In addition, in studies I, III and IV we did not have any information on serum cholesterol levels or treatment of lipid disturbance (see DAG in Appendix **Figure A.1**). We therefore tried to indirectly control for this factor by adjusting the models for the intake of fatty acids in additional sensitivity analyses, which did not affect our results.

Second, for most of the confounding variables we were able to collect precise information through the questionnaire, specifically for age, smoking, physical activity, education and BMI. For hypertension, diabetes and lipid disturbance subjects were asked to self-report, whether they had ever been treated by a doctor for one of these conditions. This could question the accuracy of the reported information. However, previous validation studies have shown acceptable agreement between self-reported information and medical records, with highest accuracy for diabetes and lowest for lipid disturbance [129, 130]. In addition, for vitamin supplement use the information reported in both questionnaires was not accurate enough. The variables were therefore categorized only into vitamin supplement users and non-users, which might have been too simplified.

Reverse causality

Temporality refers to the necessity that the cause occurs before the effect in time. This criterion must hold if we want to claim any observation to be causal [25]. In studies I to IV subjects with early disease onset during follow-up might have had other pre-clinical manifestations prior to the diagnosis of the disease and beginning of follow-up, which could have affected their dietary habits assessed at baseline. Although we excluded subjects with diagnosis of any CVD or cancer before enrolment in the study, undiagnosed conditions could still produce reverse causation of the cause and effect. Therefore, to avoid reverse causation, it can help to have large prospective studies with a long follow-up that allows exclusion of cases occurring at the early stages of follow-up related to undiagnosed health conditions [131]. When repeating the main analyses after excluding cases, which occurred during the first couple of years of follow-up, the findings in studies I to IV were not affected. Nevertheless, the exact time frame to avoid reverse causation is not known. Overall, prospective studies cannot offer definitive conclusions on cause and effect, especially if potential reverse causation remains, which could overestimate the strength of the association [132].

7.2.3 Assessment of the exposure

Food frequency questionnaire

In studies I to IV dietary intake was assessed by a self-reported and validated FFQs collected at baseline. Nevertheless, there are still some problems that can cause misclassification of the exposure.

First of all, dietary intake was assessed only once at baseline and changes in dietary habits could therefore not be captured.

Second, in our studies missing values on consumption frequencies were interpreted as null intakes [91]. However, it has previously been shown that items, which are left blank, are not consumed. In a study conducted within the Nurses' Health Study II, 64% of the blank items were reported to be consumed never or less than once per month and 20% were consumed only 1-3 times per month [92], which makes the assumption of missing values to be zero reasonable in nutritional epidemiology [92].

Third, inaccurate reporting of consumption frequencies and portion size can lead to implausibly high or low responses. This problem was specifically the case for the FFQ use in the WLHC, because of the open-ended response options for frequency values. Therefore, it is

necessary to make additional decisions regarding allowable ranges for nutrient intakes [92]. To further reduce potential exposure misclassification in our studies, all subjects who had left the main section of the FFQ blank (studies I, III and IV) or who reported extreme energy intake (studies I-IV), were excluded from the study. In addition, energy adjustment of nutrient intakes will to a large extent compensate for overall under- and over-reporting of food intake [92].

Overall, we cannot rule out that some misclassification of the exposure remains. Because of the prospective design, the proportion of misclassified subjects on the exposure variable does not depend on the outcome variable and therefore the misclassification is most likely non-differential. This would bias effect estimates of the association towards the null, which would lead to an underestimation of any positive or inverse relationship [25].

NEAC and the antioxidant food database

Strengths and limitations of using NEAC and antioxidant food databases in relation to our studies need to be discussed.

One big strength of studies I, III and IV was the comparison of two large antioxidant food databases to estimate dietary NEAC. Antioxidant values of foods might differ between countries. However, when comparing the two databases, the correlation between FRAP values was high. Further, our estimates stayed robust, indicating that the choice of the antioxidant food database has minor influence on the estimation of the association between dietary NEAC and CVD risk.

In studies I to IV we were further able to compare two different chemical assays to measure dietary NEAC, one with HAT and one with SET characteristic (i.e. FRAP and TRAP). This allowed us to control, whether our results were affected by the chemical assay used.

Limitations of the studies are that we were not able to collect biological samples from cohort participants at baseline. Previous acute dietary intervention studies have observed the ability of diet to modulate plasma NEAC [65]. However, we could not validate the concordance between plasma NEAC and FFQ derived NEAC in any of the cohorts used for our studies. The assessment of dietary NEAC through an FFQ has been evaluated previously in a similar study population and correlation coefficients between plasma NEAC and FFQ based NEAC were 0.28 for FRAP and 0.31 for TRAP [133]. Although these correlations seem low, it is important to note that plasma NEAC concentrations depend on several factors, such as

physiological differences in absorption and disposal of antioxidants, but also genetic variation affecting plasma TAC. In addition, the time of consumption of foods plays an important role, as plasma NEAC was shown to be highest after the first and second hour of consumption in acute studies, with a metabolic decrease after about four hours [65]. Nevertheless, assessing dietary NEAC through an FFQ to estimate the total dietary antioxidant capacity was shown to be a valid and useful tool in epidemiological studies [133-135].

For some food items, such as meat and dairy products, NEAC values were missing in the antioxidant food databases, especially when using the database from Pellegrini et al. in study II. However, for animal-based foods the antioxidant content is known to be low [70]. Unfortunately, we were not able to assess NEAC values for cooking oils (e.g. olive oil), because information on portion size and consumption frequency was not available in the FFQ. Olive oil contains high amounts of polyphenols, which are compounds with high antioxidant activity [136]. However, consumption of olive oil was not common in Sweden in the 1990s, which is why exclusion of olive oil might not have affected our results. Nevertheless, this measurement error could have lead to misclassification of the exposure. However, errors in nutrient intake usually lead to an underestimation of the observed estimates [25].

7.2.4 Assessment of the outcome

Another strength of these studies was the assessment of CVDs by linking individual PINs to well-managed National Health registers with nation-wide coverage and mandatory reporting from health care systems, which allowed an essentially complete follow-up of the study participants. However, the PIN can bring some pitfalls with it, which can cause incorrect identification of individuals and disease outcomes. First, PINs are sometimes coded incorrectly, which mainly arises due to incorrect recording of date of birth and sex among newborns and immigrants. Second, since PINs can be re-used, which is most commonly the case if immigrants are assigned a PIN, sometimes two individuals will end up with the same PIN. However, Statistics Sweden and the Swedish Board of Health and Welfare have systems to control for erroneous PINs, and if two individuals have the same PIN they will be assigned different serial numbers for the purpose of medical research [96].

In addition, validity of the Inpatient Register has been assessed previously, and for CVDs the validity was shown to be high. For the diagnosis of MI a positive predictive value (PPV) of at least 98% has been reported previously. For total stroke a PPV of 98.6% was reported earlier, whereas for non-fatal stroke the PPV was lower (68.5%). The validity of HF in the Swedish

Inpatient Register is lower than for other CVDs [101]. Nevertheless, when including cases with a primary diagnosis of HF, as we did in study IV, the validity increases [101] and PPVs of at least 88% were reported for the Swedish Inpatient Register [97].

Overall, remaining misclassification of the disease cannot be ruled out. However, given the prospective design of the study, the misclassification is likely non-differential, and under the assumption that the misclassification is independent from other errors, this would produce bias towards the null [25].

7.2.5 Statistical considerations

Cox proportional hazards regression model

The Cox PH regression model is the most widely used statistical model to estimate hazard ratios in epidemiological studies. A reason for its popularity is the semiparametric property of the Cox model, which does not require the investigator to make any assumption about the survival distribution of the data [106, 137]. However, there are also some limitations coming with the Cox model and hazard ratios. First of all, if the proportional hazards assumption is violated for the main exposure, the HR looses its validity. In this case, other approaches to account for the time-varying effect should be used [106]. Further limitations of the Cox model are that the baseline hazard cannot be estimated. As information on the background risk is not available, we cannot estimate absolute background risk, which would be relevant for translating the information into meaningful public health messages [138, 139]. In our study, the PH assumption was not violated for the main exposure variable. In general, as we were mainly interested in the relative effect of the exposure on the hazard rate of CVD, the Cox model was a reasonable choice in our studies.

Choice of time-scale

When using Cox PH models, there are two possible choices for the time scale, which is either time on the study, referring to calendar time from, e.g. a baseline survey, or chronological age [140]. It is not straight forward, which time scale is more appropriate to use and choice of time-scale varies largely between studies. It is important to note, that the choice of time-scale might affect the estimates. It has been suggested that if the baseline hazard is exponential or the age at entry is independent from the covariates used in the analysis, both time-scales will yield similar results [120]. However, previous simulation studies have shown, that the two time-scales can lead to different results, even if these conditions hold [141]. Although in 1997

Kom and colleagues [120] recommended to use age as time-scale, because this procedure is more meaningful and less restrictive compared to using time-on-study as the time scale, the discussion is still ongoing. Since the baseline time did not have any specific meaning in relation to the disease in our studies, we decided to use age as time-scale, as it is further an indirect way to adjust for an age effect.

Effect modification

In studies I to IV we tried to further investigate, whether the effect of dietary NEAC on CVD risk was modified by certain covariates. Indeed, the effect seemed to differ between some variables, such as smoking and supplement use. However, we did not detect any significant effect modification on the multiplicative or additive scale, when using statistical tests, which is in line with the findings reported in previous studies. Nevertheless, these findings should be interpreted with caution and some potential pitfalls need to be addressed.

As the variables of interest were dichotomized based on certain assumptions, this could have lead to residual confounding and a loss of statistical power to detect any effect [142]. Since some assumptions might have been over simplified, this could have further hidden any potential effect due to interaction. For example, in studies I, III and IV the variable smoking was regrouped into two categories only, by collapsing never and former smokers together in one category and leaving current smokers in the other. Therefore, some important information of the former smoking group might have been lost. A previous study investigated the effect of a Healthy Nordic Food Index on the risk of CVD in the WLHC and found an interaction with smoking status, reporting an inverse association among the subgroup of former smokers but no association with never or current smokers [84].

Overall, we cannot guarantee that the effect of dietary NEAC on the risk of CVD is not modified by certain lifestyle factors and this should be further investigated with careful considerations of methods and assumptions made to assess interaction.

7.2.6 Generalizability

Generalizability of our findings to other populations needs to be discussed. First of all, as studies I, III and IV were conducted only among women, these findings are not generalizable to men. Further, although the study population was randomly selected, only 51% of the initially invited women participated in the study, which might affect representativeness of the study cohort. Therefore, the generalizability to the general Swedish female population might

be questioned. Similarly, the self-selected nature of the Swedish National March Cohort might unintentionally affect the representativeness of the cohort and the generalizability of the study findings [86].

Representativeness and validity of population-based and selected cohort studies have been largely discussed in the epidemiological literature. If taken to an extreme, pursuing representativeness can defeat the goal of validly identifying causal relations [25]. Problems with population-based cohort studies are low-response rates at baseline and incomplete follow-up [143]. Therefore, selecting a study population that is representative of the larger population will often make it more difficult to achieve internal validity [25]. In addition, initial non-response of individuals limits external validity of the study, as non-response is barely randomly distributed across the population [144]. Due to the low response rate in the WLHC this might be true for studies I, III and IV. A selected-population, on the other hand, enhances feasibility, increases prevalence of exposure and completeness of the study, therefore increasing internal validity and precision of the study [86], which might be the case for the SNMC. Nevertheless, as the WLHC might have suffered from a healthy volunteer bias as discussed before, the limited validity might be balanced against a high quality of the information retrieved through the questionnaire.

7.3 Final remarks on antioxidants and NEAC

Diet plays an important role in the regulation of plasma redox status and is the main external contributor to the bodies defense against oxidative damage [65]. The role of antioxidants, together with the promising results of these and previous studies investigating the association between NEAC on the risk of CVD and the disappointing findings reported in randomized controlled trials should be discussed.

First of all, while atherosclerosis manifests clinically in middle and late adulthood, atherosclerosis has a long asymptomatic phase of development, which often begins early in life [145]. It might therefore be impossible to show the beneficial effects of antioxidant therapy over several years if the therapy is trying to reverse the results of decades of oxidative stress [146]. In fact, many studies did not assess whether the dose of the antioxidant given in the trial actually was able to decrease oxidative stress. Therefore, individuals could have either been under-treated or not treated for long enough duration to demonstrate any effect [48]. It should be remembered that the disappointing findings from clinical trials to date do

not disprove the central role of oxidative stress and antioxidants in the atherosclerotic process [146].

Moreover, other potential explanations should be considered. The biological activity and the potency of natural antioxidants might differ from synthetic antioxidants used in clinical trials [63]. Further, isolation of antioxidants might reduce their bioactivity, which is why they might not act the same way when present in whole foods [147]. In addition, levels of antioxidants used in supplements are usually much higher than found in plant-based foods [148]. This might reduce the efficacy and increase toxicity of the compound [147]. For example, Vitamin E is a well-known antioxidant that becomes a pro-oxidant if present in high concentrations [149].

Given the disappointing findings from clinical trials on antioxidant supplements, the interest of studying dietary antioxidants consumed through whole foods has been renewed [48]. Although findings from dietary intervention studies assessing the association between diets characterized by high intake of antioxidant rich foods and the risk of CVD have been promising, it is not clear what antioxidants are responsible for this beneficial effect [48]. As the human diet contains a mixture of antioxidants with different redox potentials and interactions between them, where especially the interplay between phytochemicals has been suggested to exert potent antioxidant activity [147], the NEAC assay is a relevant tool assessing these complementary mechanisms. To sum up, this makes the use of NEAC a valuable research tool to better understand the role of dietary antioxidants in the prevention of CVDs and other chronic diseases [70].

7.4 Conclusions

In the Swedish Women's Lifestyle and Health Cohort a higher baseline dietary NEAC was associated with a reduced the risk of myocardial infarction and heart failure in young to middle aged women. No association was found between dietary NEAC and the risk of stroke.

In the Swedish National March Cohort a higher baseline dietary NEAC was related to a lower risk of overall and non-fatal myocardial infarction, whereas no association was found with fatal myocardial infarction. Further, the relationship between dietary NEAC and myocardial infarction might be stronger in women.

Overall, these findings support the hypothesis that a diet with high NEAC might protect from the development of myocardial infarction and heart failure and that the beneficial effect might be exerted through interactions between antioxidants. Whether this is true for stroke needs to be further investigated. Nevertheless, it is suggested to implement high amounts of antioxidant rich foods and beverages, such as fruits, vegetables, whole grains and tea, in the daily diet to lower the burden of cardiovascular diseases.

8 FUTURE DIRECTIONS

Diet is a key lifestyle behavior and modifiable risk factor and plays a central role in the development of public health strategies for the prevention of chronic diseases. The role of diet in relation to CVDs has been widely investigated, and some important advances have been achieved over the past decades. Findings from clinical trials investigating antioxidant supplementation in relation to CVDs have questioned a potential causal relationship between antioxidants and the risk of CVD. It is, however, now widely accepted that focus should lie on foods and dietary patterns rather than single and isolated nutrients. In addition, together with the findings presented in this theses, there is growing evidence that the beneficial effects of some specific foods and dietary patterns might be truly attributed to the dietary antioxidants found in the foods and interactions between them. Nevertheless, the exact mechanisms of dietary antioxidant interactions in vivo, as well as the interplay between endogenous and exogenous antioxidants, are still poorly understood. To draw potential causal conclusions on the effect of antioxidants on the risk of CVDs, these complex mechanisms have to be further investigated. Well-designed, long-term dietary intervention trials including plasma measurements and further consideration of genetic variation and other lifestyle factors are needed to help complementing the understanding of antioxidant mechanisms in vivo.

APPENDIX

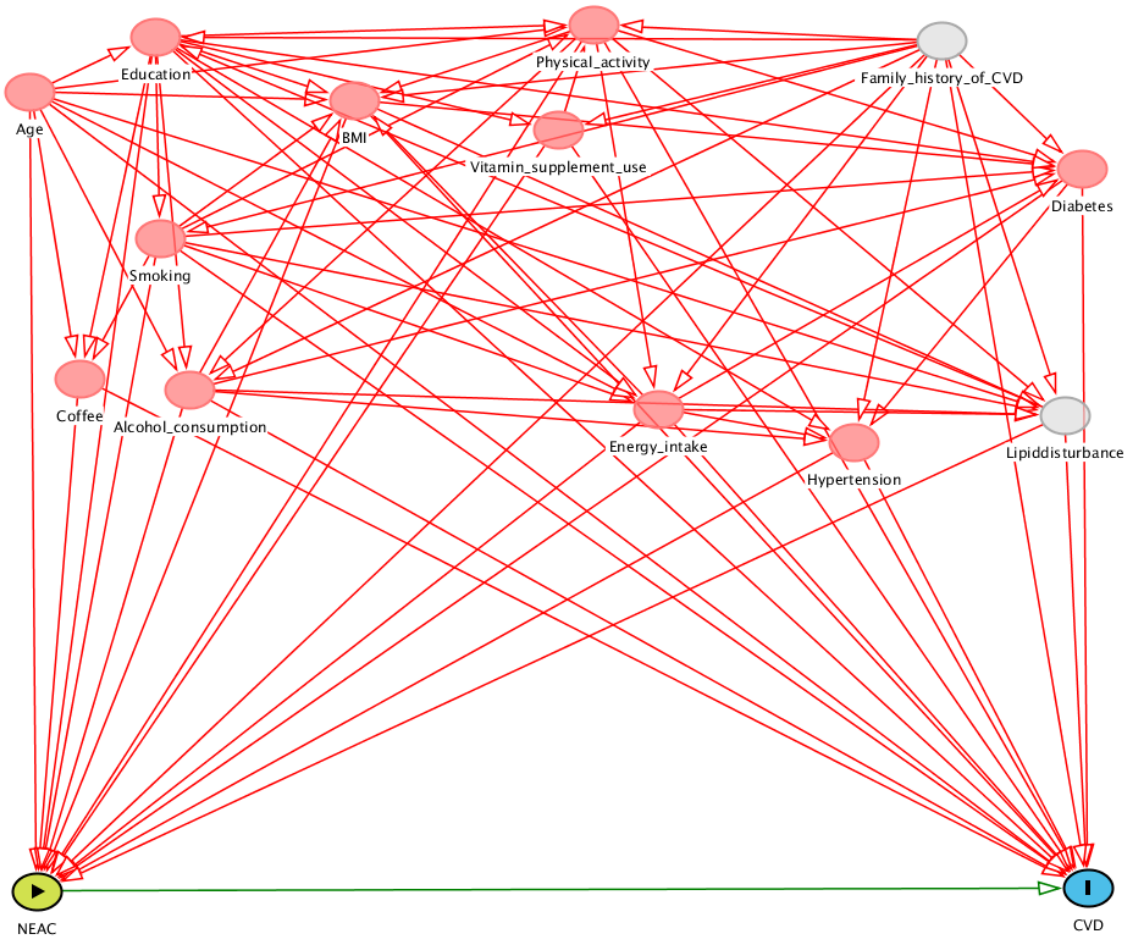

Figure A.1. Directed Acyclic Graphs for the Causal Effect of dietary NEAC and the risk of CVD in the Swedish Women's Lifestyle and Health Cohort. Yellow circe: Exposure; Blue circle: Outcome; Red circle: ancestor of the exposure and outcome; Grey circle: Unobserved variable.

ACKNOWLEDGMENTS

I have received a great deal of support and assistance throughout the years of conducting my PhD thesis.

First of all I would like to thank my supervisor **Rino Bellocco**. Thank you Rino, you have opened exciting new doors and pathways for me, you always guided, trusted and wanted the best for me, and over the past years you have become a good friend. I am deeply grateful for this interesting and challenging time in my life, which would not have been possible without you.

I would like to thank my co-supervisor **Ylva Trolle Lagerros**. Ylva, you have always believed in me, pushed and motivated me, given me amazing opportunities in the past years and supported me with great advice. You have been a wonderful mentor and a big role model and it has been an honor to work with you.

I would like to acknowledge all my co-authors. **Elisabete Weiderpass**, thank you for generously sharing the Swedish Women's Lifestyle and Health Cohort data and for the wonderful collaboration. Further, **Hans-Olov Adami** and **Olof-Nyren,** I would like to thank you for sharing the National March Cohort data and for your co-authorship. It was a great honor to work with you. You have thought me a lot with your expertise and advice in epidemiology. **Marie Löf**, it was a pleasure to work with you and thank you for providing me with great advise with all things related to nutrition. **Mauro Serafini**, thank you for introducing me to the world of antioxidants and for great advice and expertise. **Weimin Ye**, thank you for your valuable feedback.

Further, I would like to thank the Doctoral Program in Public Health for making my education possible in first place.

Next, I would like to acknowledge all my colleagues at KI and Bicocca. I have met such wonderful people throughout the past years. **Alessandra Grotta**, you are such an amazing person, statistician, mentor and friend. Thank you for all our wonderful and deep discussions, for showing me the world of statistics, for providing me your couch to sleep over, for your positivity, for believing in me. You have thought me so much, in every aspect of my life. **Francesca Ghilotti**, without you these past years would have not been the same. Thank you for always having an open ear, for always finding the right words when I needed them, for our wonderful and intense time together at Bicocca, KI and the Summer School. You have become a wonderful friend and I am forever grateful for having

had the chance to go through this experience together with you. **Daniela Mariosa**, thank you for always welcoming me at KI, for having "fika" together and for the Italian evenings. **Cecilia Damiano**, it was a pleasure to meet you, you are such a lovely person. My wonderful students **Lois Veen** and **Valentina Vitale**. It has been a pleasure to work with you. You always had the right questions and you both have thought me a lot during my time as co-supervisor. **Stephanie Bonn** and **Madeleine Wetterholm**, thank you for providing me with new challenges and a good time together. The PhD students at Bicocca, thank you for always being so welcoming and helpful in understanding the Italian system.

Finally, I would like to say a big thank you to my family and friends. **Fabian**, my partner in crime and the love of my life, thank you for being so patient with me, for always supporting me and believing in me. I know I have not been the easiest person and still you have always surprised me with your uplifting words, genuine love towards me and our family. My son **Henri**, my love and my heart, you have always showed me what the important things are in life and put my view back in the right perspective. You have made me a better person.

My mother **Virpi**, thank you for always being here for me, for always picking up the phone, for helping me out in tough times, for travelling miles and miles for us. I might finally understand you. My father **Rauli**, who is always worried about his little girl, thank you for always supporting me. My sister **Noora**, thank you for providing me with fun distraction and for being proud of your big sister.

My sisters in law, **Simone** and **Hann**a, thank you for being there and for helping me out when needed. My father in law **Roland**, thank you for always being so caring. My beloved mother in law **Christine**, you were an amazing women, mother, grandmother and friend, I truly miss you.

My wonderful friends all over the world, thank you all for your support, your words and messages, for keeping me updated what is going on outside of my office.

REFERENCES

1. Roth, G.A., et al., *Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015.* Journal of the American College of Cardiology, 2017. **70**(1): p. 1-25.

2. Townsend, N., et al., *Cardiovascular disease in Europe: epidemiological update 2016.* European heart journal, 2016. **37**(42): p. 3232-3245.

3. Mozaffarian, D., et al., *Heart disease and stroke statistics-2016 update a report from the American Heart Association.* Circulation, 2016. **133**(4): p. e38-e48.

4. Ravera, A., et al., *Nutrition and cardiovascular disease: finding the perfect recipe for cardiovascular health.* Nutrients, 2016. **8**(6): p. 363.

5. Salvayre, R., A. Negre-Salvayre, and C. Camare, *Oxidative theory of atherosclerosis and antioxidants.* Biochimie, 2015.

6. Ye, Y., J. Li, and Z. Yuan, *Effect of antioxidant vitamin supplementation on cardiovascular outcomes: a meta-analysis of randomized controlled trials.* PLoS One, 2013. **8**(2): p. e56803.

7. Aaronson, P.I., J.P. Ward, and M.J. Connolly, *The cardiovascular system at a glance*. 2012: John Wiley & Sons.

8. Benjamin, E.J., P. Muntner, and M.S. Bittencourt, *Heart disease and stroke statistics-2019 update: a report from the American Heart Association.* Circulation, 2019. **139**(10): p. e56-e528.

9. Thygesen, K., et al., *Third universal definition of myocardial infarction.* Circulation, 2012. **126**(16): p. 2020-2035.

10. Moran, A.E., et al., *The global burden of ischemic heart disease in 1990 and 2010: the Global Burden of Disease 2010 study.* Circulation, 2014. **129**(14): p. 1493-1501.

11. Mehta, L.S., et al., *Acute myocardial infarction in women: a scientific statement from the American Heart Association.* Circulation, 2016. **133**(9): p. 916-947.

12. Allen, C.L. and U. Bayraktutan, *Oxidative stress and its role in the pathogenesis of ischaemic stroke.* Int J Stroke, 2009. **4**(6): p. 461-70.

13. Gorelick, P.B., *The global burden of stroke: Persistent and disabling.* The Lancet Neurology, 2019. **18**(5): p. 417-418.

14. Sacco, R.L., et al., *An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association.* Stroke, 2013. **44**(7): p. 2064-2089.

15. Scicchitano, P., et al., *The role of endothelial dysfunction and oxidative stress in cerebrovascular diseases.* Free radical research, 2019(just-accepted): p. 1-471.

16. Schlunk, F. and S.M. Greenberg, *The pathophysiology of intracerebral hemorrhage formation and expansion.* Translational stroke research, 2015. **6**(4): p. 257-263.

17. Welty, T. and T. Horner, *Pathophysiology and treatment of subarachnoid hemorrhage.* Clinical pharmacy, 1990. **9**(1): p. 35-39.

18. Johnson, C.O., et al., *Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016.* The Lancet Neurology, 2019. **18**(5): p. 439-458.

19. Savarese, G. and L.H. Lund, *Global Public Health Burden of Heart Failure.* Card Fail Rev, 2017. **3**(1): p. 7-11.

20. Munzel, T., et al., *Pathophysiological role of oxidative stress in systolic and diastolic heart failure and its therapeutic implications.* Eur Heart J, 2015. **36**(38): p. 2555-64.

21. Bui, A.L., T.B. Horwich, and G.C. Fonarow, *Epidemiology and risk profile of heart failure.* Nat Rev Cardiol, 2011. **8**(1): p. 30-41.

22. Riedinger, M.S., et al., *Quality of life in patients with heart failure: do gender differences exist?* Heart & Lung: The Journal of Acute and Critical Care, 2001. **30**(2): p. 105-116.

23. Frohlich, J. and A. Al-Sarraf, *Cardiovascular risk and atherosclerosis prevention.* Cardiovascular Pathology, 2013. **22**(1): p. 16-18.

24. Joseph, P., et al., *Reducing the global burden of cardiovascular disease, part 1: the epidemiology and risk factors.* Circulation research, 2017. **121**(6): p. 677-694.

25. Rothman, K.J., S. Greenland, and T.L. Lash, *Modern epidemiology*. 2008: Lippincott Williams & Wilkins.

26. *Population Attributable Risk (PAR)Population attributable risk (PAR)*, in *Encyclopedia of Public Health*, W. Kirch, Editor. 2008, Springer Netherlands: Dordrecht. p. 1117-1118.

27. Yusuf, S., S. Hawken, and S. Ounpuu, *Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study (vol 364, pg 937m 2004).* Lancet, 2004. **364**(9450): p. 2020-2020.

28. O'Donnell, M.J., et al., *Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study.* The lancet, 2016. **388**(10046): p. 761-775.

29. He, J., et al., *Risk factors for congestive heart failure in US men and women: NHANES I epidemiologic follow-up study.* Archives of internal medicine, 2001. **161**(7): p. 996-1002.

30. Liguori, I., et al., *Oxidative stress, aging, and diseases.* Clinical interventions in aging, 2018. **13**: p. 757.

31. Kaliora, A.C., G.V. Dedoussis, and H. Schmidt, *Dietary antioxidants in preventing atherogenesis.* Atherosclerosis, 2006. **187**(1): p. 1-17.

32. Harrison, D., et al., *Role of oxidative stress in atherosclerosis.* The American journal of cardiology, 2003. **91**(3): p. 7-11.

33. Stoll, G. and M. Bendszus, *Inflammation and atherosclerosis: novel insights into plaque formation and destabilization.* Stroke, 2006. **37**(7): p. 1923-1932.

34. Madamanchi, N.R., A. Vendrov, and M.S. Runge, *Oxidative stress and vascular disease.* Arteriosclerosis, thrombosis, and vascular biology, 2005. **25**(1): p. 29-38.

35. Munzel, T., et al., *Impact of Oxidative Stress on the Heart and Vasculature: Part 2 of a 3-Part Series.* J Am Coll Cardiol, 2017. **70**(2): p. 212-229.

36. Wang, S., et al., *How natural dietary antioxidants in fruits, vegetables and legumes promote vascular health.* Food Research International, 2011. **44**(1): p. 14-22.

37. Kerley, C.P., *Dietary patterns and components to prevent and treat heart failure: a comprehensive review of human studies.* Nutr Res Rev, 2018: p. 1-27.

38. Hu, D., et al., *Fruits and vegetables consumption and risk of stroke: a meta-analysis of prospective cohort studies.* Stroke, 2014. **45**(6): p. 1613-9.

39. Dauchet, L., et al., *Fruit and vegetable consumption and risk of coronary heart disease: A meta-analysis of cohort studies.* Journal of Nutrition, 2006. **136**(10): p. 2588-2593.

40. Aune, D., et al., *Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality—a systematic review and dose-response meta-analysis of prospective studies.* International journal of epidemiology, 2017. **46**(3): p. 1029-1056.

41. Barrett, E.M., et al., *Whole grain, bran and cereal fibre consumption and CVD: a systematic review.* British Journal of Nutrition, 2019. **121**(8): p. 914-937.

42. Ding, M., et al., *Long-term coffee consumption and risk of cardiovascular disease: a systematic review and a dose–response meta-analysis of prospective cohort studies.* Circulation, 2014. **129**(6): p. 643-659.

43. Larsson, S.C., *Coffee, tea, and cocoa and risk of stroke.* Stroke, 2014. **45**(1): p. 309-314.

44. Fung, T.T., et al., *Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women.* Arch Intern Med, 2008. **168**(7): p. 713-20.

45. Grosso, G., et al., *A comprehensive meta-analysis on evidence of Mediterranean diet and cardiovascular disease: Are individual components equal?* Crit Rev Food Sci Nutr, 2017. **57**(15): p. 3218-3232.

46. Casas, R., et al., *Nutrition and cardiovascular health.* International journal of molecular sciences, 2018. **19**(12): p. 3988.

47. Scalbert, A., I.T. Johnson, and M. Saltmarsh, *Polyphenols: antioxidants and beyond.* The American journal of clinical nutrition, 2005. **81**(1): p. 215S-217S.

48. Leopold, J.A., *Antioxidants and coronary artery disease: from pathophysiology to preventive therapy.* Coronary artery disease, 2015. **26**(2): p. 176.

49. Gaziano, J.M., et al., *A prospective study of consumption of carotenoids in fruits and vegetables and decreased cardiovascular mortality in the elderly.* Ann Epidemiol, 1995. **5**(4): p. 255-60.

50. Tavani, A., et al., *Beta-carotene intake and risk of nonfatal acute myocardial infarction in women.* Eur J Epidemiol, 1997. **13**(6): p. 631-7.

51. Klipstein-Grobusch, K., et al., *Dietary antioxidants and risk of myocardial infarction in the elderly: the Rotterdam Study.* Am J Clin Nutr, 1999. **69**(2): p. 261-6.

52. Willcox, B.J., J.D. Curb, and B.L. Rodriguez, *Antioxidants in cardiovascular health and disease: key lessons from epidemiologic studies.* Am J Cardiol, 2008. **101**(10A): p. 75D-86D.

53. Cassidy, A., et al., *High Anthocyanin Intake Is Associated With a Reduced Risk of Myocardial Infarction in Young and Middle-Aged Women.* Circulation, 2013. **127**(2): p. 188-196.

54. Ye, Z. and H. Song, *Antioxidant vitamins intake and the risk of coronary heart disease: meta-analysis of cohort studies.* European Journal of Cardiovascular Prevention & Rehabilitation, 2008. **15**(1): p. 26-34.

55. D Archivio, M., et al., *Polyphenols, dietary sources and bioavailability.* Annali-Istituto Superiore di Sanita, 2007. **43**(4): p. 348.

56. Wang, X., et al., *Flavonoid intake and risk of CVD: a systematic review and meta-analysis of prospective cohort studies.* British Journal of Nutrition, 2014. **111**(1): p. 1-11.

57. Virtamo, J., et al., *Effect of vitamin E and beta carotene on the incidence of primary nonfatal myocardial infarction and fatal coronary heart disease.* Arch Intern Med, 1998. **158**(6): p. 668-75.

58. Lee, I.M., et al., *Vitamin E in the primary prevention of cardiovascular disease and cancer: the Women's Health Study: a randomized controlled trial.* JAMA, 2005. **294**(1): p. 56-65.

59. Sesso, H.D., et al., *Vitamins E and C in the prevention of cardiovascular disease in men: the Physicians' Health Study II randomized controlled trial.* Jama, 2008. **300**(18): p. 2123-2133.

60. Sesso, H.D., et al., *Multivitamins in the prevention of cardiovascular disease in men: the Physicians' Health Study II randomized controlled trial.* Jama, 2012. **308**(17): p. 1751-1760.

61. Lonn, E., et al., *Effects of long-term vitamin E supplementation on cardiovascular events and cancer: a randomized controlled trial.* Jama, 2005. **293**(11): p. 1338-1347.

62. Marchioli, R., et al., *Vitamin E increases the risk of developing heart failure after myocardial infarction: Results from the GISSI-Prevenzione trial.* J Cardiovasc Med (Hagerstown), 2006. **7**(5): p. 347-50.

63. Vivekananthan, D.P., et al., *Use of antioxidant vitamins for the prevention of cardiovascular disease: meta-analysis of randomised trials.* Lancet, 2003. **361**(9374): p. 2017-2023.

64. Myung, S.-K., et al., *Efficacy of vitamin and antioxidant supplements in prevention of cardiovascular disease: systematic review and meta-analysis of randomised controlled trials.* BMJ, 2013. **346**.

65. Serafini, M. and D. Del Rio, *Understanding the association between dietary antioxidants, redox status and disease: is the Total Antioxidant Capacity the right tool?* Redox Report, 2004. **9**(3): p. 145-152.

66. Prior, R.L., X. Wu, and K. Schaich, *Standardized methods for the determination of antioxidant capacity and phenolics in foods and dietary supplements.* Journal of agricultural and food chemistry, 2005. **53**(10): p. 4290-4302.

67. Pellegrini, N., et al., *Total antioxidant capacity of plant foods, beverages and oils consumed in Italy assessed by three different in vitro assays.* J Nutr, 2003. **133**(9): p. 2812-9.

68. Pellegrini, N., et al., *Total antioxidant capacity of spices, dried fruits, nuts, pulses, cereals and sweets consumed in Italy assessed by three different in vitro assays.* Mol Nutr Food Res, 2006. **50**(11): p. 1030-8.

69. Haytowitz, D.B. and S. Bhagwat, *USDA database for the oxygen radical absorbance capacity (ORAC) of selected foods, Release 2.* US Department of Agriculture, 2010: p. 10-48.

70. Carlsen, M.H., et al., *The total antioxidant content of more than 3100 foods, beverages, spices, herbs and supplements used worldwide.* Nutrition Journal, 2010. **9**.

71. Franzini, L., et al., *Food selection based on high total antioxidant capacity improves endothelial function in a low cardiovascular risk population.* Nutr Metab Cardiovasc Dis, 2012. **22**(1): p. 50-7.

72. Wang, Y., et al., *Diets high in total antioxidant capacity improve risk biomarkers of cardiovascular disease: a 9-month observational study among overweight/obese postmenopausal women.* European Journal of Nutrition, 2014. **53**(6): p. 1363-1369.

73. Kim, K., T.M. Vance, and O.K. Chun, *Greater Total Antioxidant Capacity from Diet and Supplements Is Associated with a Less Atherogenic Blood Profile in U.S. Adults.* Nutrients, 2016. **8**(1).

74. Mozaffari, H., et al., *Dietary total antioxidant capacity and cardiovascular disease risk factors: a systematic review of observational studies.* Journal of the American College of Nutrition, 2018. **37**(6): p. 533-545.

75. Rautiainen, S., et al., *Total antioxidant capacity from diet and risk of myocardial infarction: a prospective cohort of women.* Am J Med, 2012. **125**(10): p. 974-80.

76. Rossi, M., et al., *Dietary non-enzymatic antioxidant capacity and the risk of myocardial infarction: a case-control study in Italy.* Nutr Metab Cardiovasc Dis, 2014. **24**(11): p. 1246-51.

77. Rautiainen, S., et al., *Total antioxidant capacity of diet and risk of stroke: a population-based prospective cohort of women.* Stroke, 2012. **43**(2): p. 335-40.

78. Del Rio, D., et al., *Total antioxidant capacity of the diet is associated with lower risk of ischemic stroke in a large Italian cohort.* J Nutr, 2011. **141**(1): p. 118-23.

79. Colarusso, L., et al., *Dietary antioxidant capacity and risk for stroke in a prospective cohort study of Swedish men and women.* Nutrition, 2017. **33**: p. 234-239.

80. Devore, E.E., et al., *Total antioxidant capacity of the diet and major neurologic outcomes in older adults.* Neurology, 2013. **80**(10): p. 904-10.

81. Rautiainen, S., et al., *Total Antioxidant Capacity of Diet and Risk of Heart Failure: A Population-based Prospective Cohort of Women.* Am J Med, 2013. **126**(6): p. 494-500.

82. Roswall, N., et al., *Cohort Profile: The Swedish Women's Lifestyle and Health cohort.* Int J Epidemiol, 2015.

83. Löf, M., et al., *Fruit and vegetable intake and risk of cancer in the Swedish women's lifestyle and health cohort.* Cancer Causes & Control, 2011. **22**(2): p. 283-289.

84. Roswall, N., et al., *No association between adherence to the healthy Nordic food index and cardiovascular disease amongst Swedish women: a cohort study.* Journal of internal medicine, 2015. **278**(5): p. 531-541.

85. Löf, M., et al., *Prospective study of coffee consumption and all-cause, cancer, and cardiovascular mortality in Swedish women.* European journal of epidemiology, 2015. **30**(9): p. 1027-1034.

86. Trolle Lagerros, Y., et al., *Cohort profile: the Swedish National March Cohort.* International journal of epidemiology, 2016. **46**(3): p. 795-795e.

87. Hantikainen, E., et al., *Prospective study of dietary Non Enzymatic Antioxidant Capacity on the risk of hip fracture in the elderly.* Bone, 2016. **90**: p. 31-36.

88. Wolk, A., et al., *A prospective study of association of monounsaturated fat and other types of fat with risk of breast cancer.* Archives of Internal Medicine, 1998. **158**(1): p. 41-45.

89. SAMET, J.M., C.G. HUMBLE, and B.E. SKIPPER, *Alternatives in the collection and analysis of food frequency interview data.* American journal of epidemiology, 1984. **120**(4): p. 572-581.

90. Messerer, M., S.-E. Johansson, and A. Wolk, *The validity of questionnaire-based micronutrient intake estimates is increased by including dietary supplement use in Swedish men.* The Journal of nutrition, 2004. **134**(7): p. 1800-1805.

91. Michels, K.B. and W.C. Willett, *Self-administered semiquantitative food frequency questionnaires: patterns, predictors, and interpretation of omitted items.* Epidemiology, 2009. **20**(2): p. 295-301.

92. Willett, W., *Nutritional epidemiology*. Vol. 40. 2012: Oxford University Press.

93. *National Food Administration (1998) Food composition tables. Uppsala, Sweden*

.

94. Willett, W.C., G.R. Howe, and L.H. Kushi, *Adjustment for total energy intake in epidemiologic studies.* Am J Clin Nutr, 1997. **65**(4 Suppl): p. 1220S-1228S; discussion 1229S-1231S.

95. Hu, F.B., et al., *Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements.* American journal of epidemiology, 1999. **149**(6): p. 531-540.

96. Ludvigsson, J.F., et al., *The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research.* European Journal of Epidemiology, 2009. **24**(11): p. 659-667.

97. Ludvigsson, J.F., et al., *External review and validation of the Swedish national inpatient register.* BMC Public Health, 2011. **11**: p. 450.

98. Barlow, L., et al., *The completeness of the Swedish Cancer Register–a sample survey for year 1998.* Acta oncologica, 2009. **48**(1): p. 27-33.

99. Brooke, H.L., et al., *The Swedish cause of death register.* European journal of epidemiology, 2017. **32**(9): p. 765-773.

100. Ludvigsson, J.F., et al., *Registers of the Swedish total population and their use in medical research.* European journal of epidemiology, 2016. **31**(2): p. 125-136.

101. Ingelsson, E., et al., *The validity of a diagnosis of heart failure in a hospital discharge register.* Eur J Heart Fail, 2005. **7**(5): p. 787-91.

102. Lagerros, Y.T., et al., *Validity and reliability of self-reported total energy expenditure using a novel instrument.* European journal of epidemiology, 2006. **21**(3): p. 227-236.

103. Lagerros, Y.T., et al., *Measures of physical activity and their correlates: the Swedish National March Cohort.* European journal of epidemiology, 2009. **24**(4): p. 161-169.

104. Ainsworth, B.E., et al., *Compendium of physical activities: classification of energy costs of human physical activities.* Medicine and science in sports and exercise, 1993. **25**(1): p. 71-80.

105. Clark, T., et al., *Survival analysis part I: basic concepts and first analyses.* British journal of cancer, 2003. **89**(2): p. 232.

106. Kleinbaum, D.G. and M. Klein, *Evaluating the proportional hazards assumption*, in *Survival analysis*. 2012, Springer. p. 161-200.

107. Bradburn, M.J., et al., *Survival analysis part II: multivariate data analysis–an introduction to concepts and methods.* British journal of cancer, 2003. **89**(3): p. 431.

108. Bradburn, M., et al., *Survival analysis Part III: multivariate data analysis–choosing a model and assessing its adequacy and fit.* British journal of cancer, 2003. **89**(4): p. 605.

109. Schoenfeld, D., *Partial residuals for the proportional hazards regression model.* Biometrika, 1982. **69**(1): p. 239-241.

110. Desquilbet, L. and F. Mariotti, *Dose-response analyses using restricted cubic spline functions in public health research.* Statistics in medicine, 2010. **29**(9): p. 1037-1057.

111. Lambert, P.C. and P. Royston, *Further development of flexible parametric models for survival analysis.* The Stata Journal, 2009. **9**(2): p. 265-290.

112. Harrell, F., *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. 2015: Springer.

113. VanderWeele, T.J. and M.J. Knol, *A tutorial on interaction.* Epidemiologic Methods, 2014. **3**(1): p. 33-72.

114. Knol, M.J., et al., *Estimating measures of interaction on an additive scale for preventive exposures.* Eur J Epidemiol, 2011. **26**(6): p. 433-8.

115. Li, R. and L. Chambless, *Test for additive interaction in proportional hazards models.* Annals of epidemiology, 2007. **17**(3): p. 227-236.

116. Perkins, N.J., et al., *Principled approaches to missing data in epidemiologic studies.* American journal of epidemiology, 2017. **187**(3): p. 568-575.

117. StataCorp, L., *Stata multiple-imputation reference manual.* Accessed at, 2013.

118. Horton, N.J. and K.P. Kleinman, *Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models.* Am Stat, 2007. **61**(1): p. 79-90.

119. Berdanier, C.D., J.T. Dwyer, and E.B. Feldman, *Handbook of nutrition and food*. 2007: CRC press.

120. Kom, E.L., B.I. Graubard, and D. Midthune, *Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale.* American journal of epidemiology, 1997. **145**(1): p. 72-80.

121. Textor, J., et al., *Robust causal inference using directed acyclic graphs: the R package 'dagitty'.* International journal of epidemiology, 2016. **45**(6): p. 1887-1894.

122. Jann, B., *Stata tip 8: Splitting time-span records with categorical time-varying covariates.* Stata journal, 2004. **4**(199-2016-2396): p. 221-222.

123. Hallström, B.r., et al., *Stroke incidence and survival in the beginning of the 21st century in southern Sweden: comparisons with the late 20th century and projections into the future.* Stroke, 2008. **39**(1): p. 10-15.

124. Bushnell, C.D., *Stroke in women: risk and prevention throughout the lifespan.* Neurol Clin, 2008. **26**(4): p. 1161-76, xi.

125. O'Donnell, M.J., et al., *Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study.* Lancet, 2016. **388**(10046): p. 761-75.

126. de Glas, N.A., et al., *Performing survival analyses in the presence of competing risks: a clinical example in older breast cancer patients.* Journal of the National Cancer Institute, 2015. **108**(5): p. djv366.

127. Putter, H., M. Fiocco, and R.B. Geskus, *Tutorial in biostatistics: competing risks and multi-state models.* Statistics in medicine, 2007. **26**(11): p. 2389-2430.

128. Hinchliffe, S.R. and P.C. Lambert, *Extending the flexible parametric survival model for competing risks.* The Stata Journal, 2013. **13**(2): p. 344-355.

129. Margolis, K.L., et al., *Validity of diabetes self-reports in the Women's Health Initiative: comparison with medication inventories and fasting glucose measurements.* Clinical trials, 2008. **5**(3): p. 240-247.

130. Huerta, J.M., et al., *Accuracy of self-reported diabetes, hypertension, and hyperlipidemia in the adult Spanish population. DINO study findings.* Revista Española de Cardiología (English Edition), 2009. **62**(2): p. 143-152.

131. Lawlor, D.A., et al., *Reverse causality and confounding and the associations of overweight and obesity with mortality.* Obesity, 2006. **14**(12): p. 2294-2304.

132. Sattar, N. and D. Preiss, *Reverse causality in cardiovascular epidemiological research: more common than imagined?* 2017, Am Heart Assoc.

133. Rautiainen, S., et al., *The validity and reproducibility of food-frequency questionnaire-based total antioxidant capacity estimates in Swedish women.* Am J Clin Nutr, 2008. **87**(5): p. 1247-53.

134. Puchau, B., et al., *Dietary total antioxidant capacity is negatively associated with some metabolic syndrome features in healthy young adults.* Nutrition, 2010. **26**(5): p. 534-541.

135. Pellegrini, N., et al., *Development and validation of a food frequency questionnaire for the assessment of dietary total antioxidant capacity. (vol 137, pg 93, 2007).* Journal of Nutrition, 2007. **137**(6): p. 1499-1499.

136. Pellegrini, N., et al., *Direct analysis of total antioxidant activity of olive oil and studies on the influence of heating.* J Agric Food Chem, 2001. **49**(5): p. 2532-8.

137. George, B., S. Seals, and I. Aban, *Survival analysis and regression models.* Journal of Nuclear Cardiology, 2014. **21**(4): p. 686-694.

138. Royston, P. and P.C. Lambert, *Flexible parametric survival analysis using Stata: beyond the Cox model.* 2011.

139. Uno, H., et al., *Alternatives to hazard ratios for comparing efficacy or safety of therapies in noninferiority studies.* Annals of internal medicine, 2015. **163**(2): p. 127.

140. Cheung, Y.B., F. Gao, and K.S. Khoo, *Age at diagnosis and the choice of survival analysis methods in cancer epidemiology.* Journal of clinical epidemiology, 2003. **56**(1): p. 38-43.

141. Chalise, P., E. Chicken, and D. McGee, *Time scales in epidemiological analysis: an empirical comparison.* arXiv preprint arXiv:1502.02534, 2015.

142. Royston, P., D.G. Altman, and W. Sauerbrei, *Dichotomizing continuous predictors in multiple regression: a bad idea.* Statistics in medicine, 2006. **25**(1): p. 127-141.

143. Nohr, E.A. and J. Olsen, *Commentary: Epidemiologists have debated representativeness for more than 40 years--has the time come to move on?* Int J Epidemiol, 2013. **42**(4): p. 1016-7.

144. Szklo, M., *Population-based cohort studies.* Epidemiol Rev, 1998. **20**(1): p. 81-90.

145. Hong, Y.M., *Atherosclerotic cardiovascular disease beginning in childhood.* Korean circulation journal, 2010. **40**(1): p. 1-9.

146. Steinhubl, S.R., *Why have antioxidants failed in clinical trials?* The American journal of cardiology, 2008. **101**(10): p. S14-S19.

147. Liu, R.H., *Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals.* Am J Clin Nutr, 2003. **78**(3 Suppl): p. 517S-520S.

148. Miller, E.R., et al., *Meta-analysis: High-dosage vitamin E supplementation may increase all-cause mortality.* Annals of Internal Medicine, 2005. **142**(1): p. 37-46.

149. Carocho, M. and I.C. Ferreira, *A review on antioxidants, prooxidants and related controversy: natural and synthetic compounds, screening and analysis methodologies and future perspectives.* Food Chem Toxicol, 2013. **51**: p. 15-25.